```
                        GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      October  6, 2004, 15:50:37 ; Search time 123 Seconds
                                          (without alignments)
                                          399.701 Million cell updates/sec

Title:         US-10-009-792C-19
Perfect score: 896
Sequence:    1 TPLGPASSLPQSFLLKCLEQ.........SHLQSFLEVSYRVLRHLAQP 174

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:    1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:        1586107


Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_29Jan04:*
             1: geneseqp1980s:*
             2: geneseqp1990s:*
             3: geneseqp2000s:*
             4: geneseqp2001s:*
             5: geneseqp2002s:*
             6: geneseqp2003as:*
             7: geneseqp2003bs:*
             8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
          score greater than or equal to the score of the result being printed,
          and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	896	100.0	174	1	AAP80965	Aap80965 G-CSF ded
2	896	100.0	174	1	AAP90470	Aap90470 G-colony
3	896	100.0	174	2	AAR13679	Aar13679 Granulocy
4	896	100.0	174	2	AAR26909	Aar26909 Human G-C
5	896	100.0	174	2	AAR34707	Aar34707 Human CSF
6	896	100.0	174	2	AAW02204	Aaw02204 Human G-C
7	896	100.0	174	3	AAY78937	Aay78937 Granulocy
8	896	100.0	174	3	AAB14850	Aab14850 Human gra
9	896	100.0	174	3	AAY97016	Aay97016 Mature gr
10	896	100.0	174	4	AAE12153	Aae12153 Human hG-
11	896	100.0	174	4	AAB61930	Aab61930 Human gra
12	896	100.0	174	4	AAB85332	Aab85332 Codon opt
13	896	100.0	174	4	AAM52110	Aam52110 Human G-C
14	896	100.0	174	4	AAE11983	Aae11983 Human wil
15	896	100.0	174	5	AAU79877	Aau79877 Human gra
16	896	100.0	174	5	AAU97116	Aau97116 Wild-type
17	896	100.0	174	5	AAE14694	Aae14694 Human gra
18	896	100.0	174	6	ABR55839	Abr55839 Human gra
19	896	100.0	174	6	AAE30598	Aae30598 Human gra
20	896	100.0	174	6	ABG74368	Abg74368 Partial h
21	896	100.0	174	6	ABR62687	Abr62687 Human gra
22	896	100.0	174	7	ADC56722	Adc56722 Human pro
23	896	100.0	174	7	ABR61548	Abr61548 Human gra
24	896	100.0	174	7	ABR61549	Abr61549 Human gra
25	896	100.0	175	1	AAP71030	Aap71030 Sequence
26	896	100.0	175	1	AAP70732	Aap70732 Sequence
27	896	100.0	175	1	AAP90107	Aap90107 Human gra
28	896	100.0	175	1	AAP90170	Aap90170 Human gra
29	896	100.0	175	1	AAP91070	Aap91070 Human gra
30	896	100.0	175	2	AAW08486	Aaw08486 Platelet
31	896	100.0	175	2	AAR98465	Aar98465 Human gra
32	896	100.0	175	2	AAW07533	Aaw07533 Recombina
33	896	100.0	175	2	AAR94394	Aar94394 PEGylated
34	896	100.0	175	2	AAW84297	Aaw84297 Human gra
35	896	100.0	175	2	AAY78936	Aay78936 Granulocy
36	896	100.0	175	3	AAB14851	Aab14851 Human gra
37	896	100.0	175	3	AAY97017	Aay97017 Mature gr
38	896	100.0	175	3	AAB23758	Aab23758 Human gra
39	896	100.0	175	4	AAG67563	Aag67563 Amino aci
40	896	100.0	175	4	AAG67502	Aag67502 Amino aci
41	896	100.0	175	4	AAE12154	Aae12154 Human hG-
42	896	100.0	175	4	AAM52109	Aam52109 Human G-C
43	896	100.0	175	4	AAM51536	Aam51536 Recombina
44	896	100.0	175	5	AAO19311	Aao19311 Branched
45	896	100.0	175	6	ABP97386	Abp97386 Recombina

```
                              ALIGNMENTS

RESULT 1
AAP80965
ID   AAP80965 standard; protein; 174 AA.
XX
AC   AAP80965;
XX
DT   15-NOV-1990  (first entry)
XX
DE   G-CSF deduced from gene isolate from peripheral blood macrophages.
XX
KW   Granulocyte colony stimulating factor; hG-CSF; macrophage;
KW   tumour therapy; leukaemia.
XX
OS   Homo sapiens.
XX
PN   EP272703-A.
XX
PD   29-JUN-1988.
XX
PF   23-DEC-1987;    87EP-00119157.
XX
PR   23-DEC-1986;    86JP-00306799.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Kuga T, Komatsu Y, Miyaji H, Sato M, Okabe M, Morimoto M;
PI   Itoh S, Yamasaki M;
XX
DR   WPI; 1988-176825/26.
DR   N-PSDB; AAN80947.
XX
PT   Human granulocyte colony stimulating factor polypeptide derivs. - having
PT   at least one different aminoacid, giving high specific activity and
PT   stability.
XX
PS   Disclosure; Page 2; 68pp; English.
XX
CC   The sequence of the gene is carried on plasmid pCSF1-2 isolated from a
CC   cDNA library prepared from mRNA extracted from peripheral blood
CC   macrophages. The deduced protein sequence agrees with those determined
CC   for G-CSF from the human squamous cell line CHU-II and the human bladder
CC   cancer cell line 5637. The DNA can be used as a "master gene" for the
CC   construction of mutant genes which encode variants of hG-CSF which differ
CC   by at least one AA. See also AAP80963 and AAP80964, and AAP82874-92
XX
SQ   Sequence 174 AA;

     Query Match          100.0%;  Score 896;  DB 1;  Length 174;
```

Left column

Best Local Similarity   100.0%;   Pred. No. 3.8e-90;
Matches   174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 2
AAP90470
ID   AAP90470 standard; protein; 174 AA.
XX
AC   AAP90470;
XX
DT   25-MAR-2003   (revised)
DT   01-NOV-1989   (first entry)
XX
DE   G-colony stimulating factor.
XX
KW   G-colony stimulating factor; G-CSF; lysine-depleted variant;
KW   site-directed mutagenesis; human.
XX
OS   Homo sapiens.
XX
PN   WO8905824-A.
XX
PD   29-JUN-1989.
XX
PF   22-DEC-1988;   88WO-US004633.
XX
PR   23-DEC-1987;   87US-00137043.
XX
PA   (GEMY) GENETICS INST INC.
XX
PI   Shaw G;
XX
DR   WPI; 1989-206594/28.
DR   N-PSDB; AAN90256.
XX
PT   New lysine depleted variants of polypeptide - opt. modified with
PT   hydrophilic residues, biologically active but with altered solubility,
PT   stability, etc.
XX
PS   Disclosure; Fig 4; 35pp; English.
XX
CC   DNA of G-CSF (G-colony stimulating factor) (see corresp. AAN90256). Used
CC   in the patent to create lysine depleted variants by site-directed
CC   mutagenesis, or synthesis. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 174 AA;

Query Match   100.0%;   Score 896;   DB 1;   Length 174;
Best Local Similarity   100.0%;   Pred. No. 3.8e-90;
Matches   174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

Right column

```
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 3
AAR13679
ID   AAR13679 standard; protein; 174 AA.
XX
AC   AAR13679;
XX
DT   25-MAR-2003   (revised)
DT   13-NOV-1991   (first entry)
XX
DE   Granulocyte colony stimulating factor.
XX
KW   G-CSF; neutropaenia.
XX
OS   Synthetic.
XX
PN   GB2241505-A.
XX
PD   04-SEP-1991.
XX
PF   27-FEB-1991;   91GB-00041260.
XX
PR   27-FEB-1990;   90GB-00004390.
PR   27-FEB-1991;   91GB-00004126.
XX
PA   (ICIL) IMPERIAL CHEM IND PLC.
XX
PI   Hockney RC, Kara BV;
XX
DR   WPI; 1991-262013/36.
DR   N-PSDB; AAQ13329.
XX
PT   Polypeptide prodn. by fermenting transformed host cells - in presence of
PT   non-auxotrophic aminoacid to increase prod. accumulation, esp. for
PT   granulocyte colony stimulating factor.
XX
PS   Disclosure; Fig 6; 69pp; English.
XX
CC   The protein was expressed from a synthetic gene designed to include
CC   several restriction sites to facilitate manipulation and to have A/T rich
CC   codons at the 5' end of the coding region. Other codons were chosen as
CC   those preferred for expression in E.coli. The gene was cloned into an
CC   expression plasmid for the prodn of G-CSF for use in the management of
CC   neuropaenia etc. See also AAR14224. (Updated on 25-MAR-2003 to correct PR
CC   field.)
XX
SQ   Sequence 174 AA;

Query Match   100.0%;   Score 896;   DB 2;   Length 174;
Best Local Similarity   100.0%;   Pred. No. 3.8e-90;
Matches   174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 4
AAR26909
ID   AAR26909 standard; protein; 174 AA.
XX
AC   AAR26909;
XX

```
DT   25-MAR-2003 (revised)
DT   20-MAY-1998 (first entry)
XX
DE   Human G-CSF.
XX
KW   G-CSF; granulocyte colony stimulating factor; inducible selection;
KW   ricin A; immunotoxins; neutropaenia; chemotherapy; radiation; therapy;
KW   autologous bone marrow transplantation; AIDS; severe infection;
KW   myelodysplastic syndromes; abnormal granulocyte function.
XX
OS   Synthetic.
XX
PN   EP502637-A2.
XX
PD   09-SEP-1992.
XX
PF   21-FEB-1992; 92EP-00301465.
XX
PR   26-FEB-1991; 91GB-00004017.
PR   29-APR-1991; 91GB-00009188.
XX
PA   (ICIL ) IMPERIAL CHEM IND PLC.
PA   (ZENE ) ZENECA LTD.
XX
PI   Barth PT;
XX
DR   WPI; 1992-301941/37.
DR   N-PSDB; AAQ28371.
XX
PT   Vector contg. inducible selection gene - encodes heterologous polypeptide
PT   e.g. ricin A and G-CSF.
XX
PS   Claim 8; Fig 10; 73pp; English.
XX
CC   This sequence is encoded by the synthetic gene AAQ28371. The plasmid
CC   containing the coding gene contains a novel inducible selection gene
CC   which is only expressed during the construction and testing phases of
CC   genetic manipulation. When the subsequent plasmid carrying the cloned
CC   gene is stably maintained in its bacterial host, the need for selection
CC   ceases. Cultures grown to express the cloned gene prod. will not require
CC   addn. of the selection drug and will not express the prod. of the
CC   selection gene. In this case the product of the cloned gene is human G-
CC   CSF which is used to treat neuropaenia associated with chemotherapy,
CC   radiation therapy, or autologous bone marrow transplantation, to
CC   stimulate bone marrow suppression associated with AIDS, in the treatment
CC   of myelodysplastic syndromes characterised by granulocyte functional
CC   abnormalities and as an adjunct to the treatment of severe infections.
CC   (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to
CC   correct PA field.)
XX
SQ   Sequence 174 AA;

Query Match             100.0%;  Score 896;  DB 2;  Length 174;
Best Local Similarity   100.0%;  Pred. No. 3.8e-90;
Matches  174;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 5
AAR34707
ID   AAR34707 standard; protein; 174 AA.
XX
AC   AAR34707;
XX
DT   10-JUN-1993 (first entry)
XX
DE   Human CSF analogue.
XX
KW   Colony stimulating factor; cancer; tumour; sarcoma; lymphoma; melanoma;
KW   lung.
XX
OS   Homo sapiens.
XX
PN   JP05009131-A.
XX
PD   19-JAN-1993.
XX
PF   17-SEP-1991; 91JP-00236260.
XX
PR   17-SEP-1990; 90JP-00246488.
PR   25-OCT-1990; 90JP-00287707.
XX
PA   (CHUS ) CHUGAI PHARM CO LTD.
XX
DR   WPI; 1993-061608/08.
XX
PT   Anticancer drug showing good apothanaisia effect - contains
PT   immuno:therapeutic agent for cancer, human colony stimulating factor and
PT   anticancer drug.
XX
PS   Disclosure; Page 3; 10pp; Japanese.
XX
CC   The sequence is that of human colony stimulating factor, which can be
CC   with an immunotherapeutic agent for cancer and an anticancer drug in a
CC   novel anticancer compsn. The compsn. has excellent anticancer effects,
CC   not shown by sole administration of the single components. It shows good
CC   life prolonging effects and has tumour growth inhibiting effects for
CC   mammalian cancers. It is esp. effective for treatment of early stage
CC   cancer and after cancer surgery. Treatable diseases include various kinds
CC   of sarcoma, malignant lymphoma, malignant melanoma, malignant chlorionic
CC   tumour, phalia cancer, various kinds of digestive organ cancer, lung
CC   cancer; etc. See also AAR32258
XX
SQ   Sequence 174 AA;

Query Match             100.0%;  Score 896;  DB 2;  Length 174;
Best Local Similarity   100.0%;  Pred. No. 3.8e-90;
Matches  174;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 6
AAW02204
ID   AAW02204 standard; protein; 174 AA.
XX
AC   AAW02204;
XX
DT   10-OCT-1996 (first entry)
XX
DE   Human G-CSF for stimulating peripheral blood cell prodn.
XX
KW   Human; interleukin; hIL-11; granulocyte colony stimulating factor;
KW   hG-CSF; in vivo; increase; production; peripheral blood cells; hPBS;
KW   hIL-1; hIL-3; stem cell factor; leukaemia inhibitory factor;
```

```
KW   erythropoietin; GM-CSF; M-CSF; transplant; treatment; transfusion;
KW   blood system disease.
XX
OS   Homo sapiens.
XX
PN   JP08127539-A.
XX
PD   21-MAY-1996.
XX
PF   31-OCT-1994;   94JP-00266826.
XX
PR   31-OCT-1994;   94JP-00266826.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 1996-295521/30.
XX
PT   Agent contg. human interleukin 11 for increasing peripheral blood stem
PT   cell prodn. - pref. comprises human G-CSF, and is useful for peripheral
PT   blood stem cell transplants to treat blood system diseases.
XX
PS   Example 1; Page 6; 6pp; Japanese.
XX
CC   An agent contg. human interleukin 11 (hIL-11), and pref. human
CC   granulocyte colony stimulating factor (hG-CSF), i.e. the present protein,
CC   can be administered in vivo to increase the prodn. of human peripheral
CC   blood (hPBS) cells. The agent may further comprise hIL-1, hIL-3, stem
CC   cell factor, leukaemia inhibitory factor, erythropoietin, GM-CSF and M-
CC   CSF, which synergistically increase hPBS cell prodn. The amt. of hIL-11
CC   or hIL-11 plus hG-CSF is generally 0.001-100, pref. 0.1-1.0 wt. %. The
CC   agent may contain a stabiliser (e.g. serum albumin) and fillers
CC   (mannitol). The dose is 0.1-1000 microg/kg/day, and the agent can be used
CC   in hPBS cell transplants for the treatment of blood system diseases
XX
SQ   Sequence 174 AA;
XX

Query Match           100.0%;     Score 896;  DB 2;  Length 174;
Best Local Similarity 100.0%;     Pred. No. 3.8e-90;
Matches 174;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy     61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy    121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 7
AAY78937
ID   AAY78937 standard; protein; 174 AA.
XX
AC   AAY78937;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Granulocyte colony-stimulation factor (G-CSF) peptide sequence #2.
XX
KW   Granulocyte colony-stimulating factor; G-CSF; powder preparation;
KW   polymeric drug administration; mucus membrane.
XX
OS   Homo sapiens.
XX
PN   WO200002574-A1.
XX
PD   20-JAN-2000.
XX
PF   01-JUL-1999;  99WO-JP003563.
XX
PR   08-JUL-1998;   98JP-00192722.
PR   25-MAR-1999;   99JP-00081549.
XX
PA   (KIRI ) KIRIN AMGEN INC.
XX
PI   Nomura H, Ueki Y;
XX
DR   WPI; 2000-182173/16.
XX
PT   Powder preparation for mucosal administration of polymeric drug, e.g.
PT   peptide, protein, antibody, vaccine or antigen.
XX
PS   Disclosure; Page 39-40; 45pp; Japanese.
XX
CC   This sequence represents a human granulocyte colony-stimulating factor (G
CC   -CSF) polypeptide. G-CSF is used in a powder preparation which is
CC   administered via the mucus membrane. The preparation comprises a
CC   polymeric medicine and a cationic polymer. The preparation is used for
CC   the mucosal administration of polymeric pharmaceuticals and has good
CC   absorption through the mucous membrane and improved bioavailability
XX
SQ   Sequence 174 AA;
XX

Query Match           100.0%;     Score 896;  DB 3;  Length 174;
Best Local Similarity 100.0%;     Pred. No. 3.8e-90;
Matches 174;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy     61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy    121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 8
AAB14850
ID   AAB14850 standard; protein; 174 AA.
XX
AC   AAB14850;
XX
DT   19-DEC-2000  (first entry)
XX
DE   Human granulocyte colony stimulating factor #1.
XX
KW   Human; granulocyte colony stimulating factor; G-CSF; neutropaenia;
KW   bone marrow suppression; infection.
XX
OS   Homo sapiens.
XX
PN   WO200040728-A1.
XX
PD   13-JUL-2000.
XX
PF   06-JAN-2000;  2000WO-US000300.
XX
PR   06-JAN-1999;   99US-0115131P.
PR   05-FEB-1999;   99US-0118831P.
XX
PA   (XENC-) XENCOR INC.
XX
PI   Dahiyat B, Luo P;
XX
DR   WPI; 2000-465988/40.
DR   N-PSDB; AAA73324.
XX
PT   A non-naturally occurring granulopoietic activity protein (GPA) for
```

```
PT    treating granulocyte colony stimulating factor (G-CSF)-responsive disease
PT    comprises an amino acid sequence less than 95% identical to hG-CSF.
XX
PS    Disclosure; Fig 1; 63pp; English.
XX
CC    The present sequence is the protein sequence for the human granulocyte
CC    colony stimulating factor (G-CSF). This protein is involved in the
CC    proliferation and differentiation of granulocytes in the blood. The
CC    sequence was used to create the proteins of the invention, which are
CC    designated granulopoietic activity (GPA) proteins. These can be used
CC    instead of G-CSF, which has a short half-life in the blood and is
CC    unstable in storage, in treatments for neutropaenia associated with
CC    cancer therapies, radiation accidents, bone marrow transplantation, bone
CC    marrow suppression conditions such as AIDS, myelodysplastic syndromes
CC    characterised by granulocyte functional abnormalities, and severe
CC    infections. They can also be used to enhance peripheral blood progenitor
CC    cell collection
XX
SQ    Sequence 174 AA;

Query Match             100.0%;  Score 896;  DB 3;  Length 174;
Best Local Similarity   100.0%;  Pred. No. 3.8e-90;
Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy         61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy        121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 9
AAY97016
ID    AAY97016 standard; protein; 174 AA.
XX
AC    AAY97016;
XX
DT    31-OCT-2000  (first entry)
XX
DE    Mature granulocyte colony stimulating factor.
XX
KW    G-CSF; granulocyte colony stimulating factor; sustained-release;
KW    biocompatible polyol/oil suspension; anti-inflammatory.
XX
OS    Homo sapiens.
XX
PN    WO200038652-A1.
XX
PD    06-JUL-2000.
XX
PF    20-DEC-1999;  99WO-US030527.
XX
PR    23-DEC-1998;  98US-00221181.
PR    23-NOV-1999;  99US-0448205.
XX
PA    (AMGE-) AMGEN INC.
XX
PI    Goldenberg MS,  Shan D,  Beekman AC;
XX
DR    WPI; 2000-452289/39.
XX
PT    Pharmaceutical composition for the sustained-release of a biologically
PT    active agent (BAA), such as granulocyte-colony stimulating factor,
PT    comprises incorporating the BAA into a biocompatible polyol/oil
PT    suspension.
XX
PS    Claim 7; Page 11; 38pp; English.

XX
CC    A pharmaceutical composition comprising a biologically active agent (BAA)
CC    incorporated into a biocompatible polyol/oil suspension which contains a
CC    thickener is new. The compositions are used for the sustained-release of
CC    a BAA such as an interferon consensus, EPO, granulocyte-colony
CC    stimulating factor, stem cell factor, leptin, tumor necrosis factor-
CC    binding protein, interleukin-1 receptor antagonist, brain derived
CC    neurotrophic factor, glial derived neurotrophic factor, neurotrophic factor
CC    3, osteoprotegerin, granulocyte macrophage colony stimulating factor,
CC    megakaryocyte derived growth factor, keratinocyte growth factor,
CC    thrombopoietin, or novel erythropoiesis stimulating protein (claimed).
CC    The release of a medicament can be controlled to provide longer periods
CC    of consistent release that previous methods of treatment do not achieve,
CC    such as repeated injections. Blood levels of the active ingredient can be
CC    controlled, providing an enhanced prophylactic, therapeutic, or
CC    diagnostic effect as well as greater safety, patient convenience and
CC    patient compliance. The compositions can lead to dose sparing and a lower
CC    cost of protein production. Bioavailability and protein protection,
CC    stability and potency are increased
XX
SQ    Sequence 174 AA;

Query Match             100.0%;  Score 896;  DB 3;  Length 174;
Best Local Similarity   100.0%;  Pred. No. 3.8e-90;
Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy         61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy        121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 10
AAE12153
ID    AAE12153 standard; protein; 174 AA.
XX
AC    AAE12153;
XX
DT    18-DEC-2001  (first entry)
XX
DE    Human hG-CSF inserted into plasmid p19CSFm.
XX
KW    Human; granulocyte colony stimulating factor; hG-CSF; protease;
KW    Factor Xa; kanamycin resistance; endoxylanase signal peptide.
XX
OS    Homo sapiens.
XX
PN    WO200173081-A1.
XX
PD    04-OCT-2001.
XX
PF    31-MAR-2001;  2001WO-KR000549.
XX
PR    31-MAR-2000;  2000KR-00017052.
XX
PA    (KOAD ) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI    Lee S,  Jeong K;
XX
DR    WPI; 2001-616523/71.
DR    N-PSDB; AAD19771.
XX
PT    Recombinant plasmid vector comprising an endoxylanase signal sequence,
PT    human granulocyte colony stimulating factor gene and other components,
PT    when transformed into microorganism useful for preparing the colony
PT    stimulating factor.
```

```
XX
PS   Example 1; Fig 3; 50pp; English.
XX
CC   The invention relates to an Escherichia coli producing and secreting
CC   human granulocyte colony stimulating factor (hG-CSF), more specifically,
CC   to a recombinant plasmid constructed to express secretory hG-CSF in E.
CC   coli, an E. coli transformed with that plasmid to secrete hG-CSF, and a
CC   process for preparing hG-CSF using the transformed hG-CSF. The
CC   recombinant plasmid vector comprises of a kanamycin resistance gene, a
CC   promoter, an endoxylanase signal sequence, a nucleotide sequence coding
CC   for an oligopeptide consisting of 13 amino acids including 6 consecutive
CC   histidine residues and a hG-CSF. E.coli transformed with recombinant
CC   plasmid vector is useful for preparing hG-CSF. The method comprises
CC   culturing the microorganism to obtain a hG-CSF fusion protein and
CC   treating the fusion protein with a protease preferably Factor Xa, to
CC   obtain a hG-CSF, where the fusion protein is obtained from the culture by
CC   employing Ni-column. The present sequence is human hG-CSF inserted into
CC   plasmid p19CSFm
XX
SQ   Sequence 174 AA;

Query Match           100.0%;     Score 896;  DB 4;   Length 174;
Best Local Similarity 100.0%;     Pred. No. 3.8e-90;
Matches 174;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174


RESULT 11
AAB61930
ID   AAB61930 standard; protein; 174 AA.
XX
AC   AAB61930;
XX
DT   08-MAY-2001  (first entry)
XX
DE   Human granulocyte-colony stimulating factor (hG-CSF).
XX
KW   Granulocyte-colony stimulating factor; G-CSF; human; variant.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   Misc-difference 1
FT                  /note= "can be replaced by Ser"
FT   Misc-difference 2
FT                  /note= "can be replaced with met"
FT   Misc-difference 3
FT                  /note= "can be replaced with Val"
FT   Misc-difference 17
FT                  /note= "can be replaced with Ser, Thr, Ala or Gly"
XX
PN   WO200104329-A1.
XX
PD   18-JAN-2001.
XX
PF   07-JUL-2000; 2000WO-KR000733.
XX
PR   08-JUL-1999;  99KR-00027418.
XX
PA   (HANM-) HANMI PHARM CO LTD.
XX
PI   Kwon SC, Jung SY, Bae SM, Lee GS;
XX
DR   WPI; 2001-138357/14.
DR   N-PSDB; AAF56684.
XX
PT   Modified human granulocyte-colony stimulating factor (hG-CSF), retaining
PT   the biological activity of wild type hG-CSF, comprises replacement amino
PT   acids at the first, second, third and seventeenth positions.
XX
PS   Claim 1; Fig 1; 69pp; English.
XX
CC   The invention relates to a modified human granulocyte-colony stimulating
CC   factor (hG-CSF), characterized in that at least one of the first, second,
CC   third and seventeenth amino acids of wild-type hG-CSF (AAB61930) is
CC   replaced by other amino acids. The modified hG-CSF retains the biological
CC   activity of the wild-type protein. It does not have a methionine residue
CC   at the N-terminus and so can be efficiently expressed and secreted by a
CC   microorganism, when an appropriate secretory signal peptide is employed.
CC   The modified hG-CSF is useful for processes requiring hG-CSF. The present
CC   sequence represents the wild-type human hG-CSF mature protein
XX
SQ   Sequence 174 AA;

Query Match           100.0%;     Score 896;  DB 4;   Length 174;
Best Local Similarity 100.0%;     Pred. No. 3.8e-90;
Matches 174;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174


RESULT 12
AAB85332
ID   AAB85332 standard; protein; 174 AA.
XX
AC   AAB85332;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Codon optimised mature hG-CSF protein.
XX
KW   Granulocyte-colony stimulating factor; G-CSF; hemotopoietic; AIDS;
KW   chemotherapy; immunodeficiency disease; anti-HIV; cytostatic;
KW   antibacterial; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200151510-A2.
XX
PD   19-JUL-2001.
XX
PF   09-JAN-2001; 2001WO-DK000011.
XX
PR   10-JAN-2000;  2000DK-00000024.
PR   02-MAR-2000;  2000DK-00000341.
PR   16-JUN-2000;  2000DK-00000943.
XX
PA   (MAXY-) MAXYGEN APS.
XX
PI   Nissen TL, Andersen KV, Hansen CK, Mikkelsen JM, Schambye HT;
XX
DR   WPI; 2001-451839/48.
DR   N-PSDB; AAH22902.
XX
PT   Novel polypeptides exhibiting granulocyte-colony stimulating factor
```

```
PT  activity and conjugates between the polypeptide and a non-polypeptide
PT  moiety, useful for treating leukopenia, AIDS and bacterial or other
PT  infections.
XX
PS  Example 1; Page 91-92; 94pp; English.
XX
CC  The invention provides a polypeptide (PP) conjugate (I) exhibiting
CC  granulocyte-colony stimulating factor (G-CSF) activity, comprising PP
CC  with an amino acid (aa) sequence that differs from aa sequence of human G
CC  -CSF by at least one aa residue comprising an attachment group for a non-
CC  PP group and having a non-PP group bound to attachment of PP, and the
CC  polypeptide (II) portion of (I) exhibiting G-CSF activity. (I) and (II)
CC  are useful for treating general hemotopoietic disorders, including
CC  disorders arising from radiation therapy or from chemotherapy, AIDS or
CC  immunodeficiency diseases, leukopenia and bacterial or other infections.
CC  (II) is useful for therapeutic, diagnostic and other purposes and in
CC  particular finds use as intermediate products for preparation of (I). (I)
CC  and (II) are used for preventing infection in cancer patients undergoing
CC  radiation therapy, chemotherapy and bone marrow transplantations, to
CC  mobilize progenitor cells for collection in peripheral blood progenitor
CC  cell transplantations and to support treatment of patients with acute
CC  myeloid leukemia. A nucleotide sequence encoding (II) is useful in gene
CC  therapy applications. (I) has increased functional in vivo half-life,
CC  increased serum half-life, reduced renal clearance, reduced receptor-
CC  mediated clearance, reduced side effects, such as bone pain, reduced
CC  immunogenicity and/or increased bioavailability. The present sequence
CC  represents a mature hG-CSF with codon usage optimised for the encoding
CC  DNA, for expression in E. coli. This is used in the construction of
CC  synthetic genes encoding hG-CSF
XX
SQ  Sequence 174 AA;
XX

Query Match     100.0%;    Score 896;   DB 4;    Length 174;
Best Local Similarity   100.0%;   Pred. No. 3.8e-90;
Matches 174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 13
AAM52110
ID  AAM52110 standard; protein; 174 AA.
XX
AC  AAM52110;
XX
DT  28-JAN-2002  (first entry)
XX
DE  Human G-CSF polypeptide SEQ ID NO 2.
XX
KW  Human; G-CSF.
XX
OS  Homo sapiens.
XX
PN  WO200174397-A1.
XX
PD  11-OCT-2001.
XX
PF  28-MAR-2001; 2001WO-JP002555.
XX
PR  31-MAR-2000; 2000JP-00099213.
XX
PA  (KIRI ) KIRIN BEER KK.
XX
PI  Nomura H, Ueki Y;
XX
DR  WPI; 2001-662952/76.
XX
PT  Stable powdered preparation for transmucosal administration comprises
PT  polymeric drug, cationic polymer and basic amino acid.
XX
PS  Disclosure; Page 23-24; 28pp; Japanese.
XX
CC  The invention relates to a stable powdered preparation for transmucosal
CC  administration (preferably nasally) comprising a polymeric form of drug
CC  (preferably peptides), a cationic polymer and at least one basic amino
CC  acid or its salt. The present sequence is that of human G-CSF polypeptide
XX
SQ  Sequence 174 AA;
XX

Query Match     100.0%;    Score 896;   DB 4;    Length 174;
Best Local Similarity   100.0%;   Pred. No. 3.8e-90;
Matches 174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 14
AAE11983
ID  AAE11983 standard; protein; 174 AA.
XX
AC  AAE11983;
XX
DT  18-DEC-2001  (first entry)
XX
DE  Human wild-type granulocyte colony stimulating factor (G-CSF).
XX
KW  Human; interferon-gamma; IFNG; interferon-beta; interferon alpha; FSH;
KW  follicle stimulating hormone; granulocyte colony stimulating factor;
KW  G-CSF; interleukin 10; IL-10; medicament; pulmonary administration;
KW  circulatory administration.
XX
OS  Homo sapiens.
XX
PN  WO200168141-A2.
XX
PD  20-SEP-2001.
XX
DR  WPI; 2001-616274/71.
XX
PF  16-MAR-2001; 2001WO-DK000182.
XX
PR  17-MAR-2000; 2000DK-00000447.
PR  20-MAR-2000; 2000US-0190844P.
PR  25-AUG-2000; 2000WO-DK000471.
PR  13-NOV-2000; 2000WO-DK000631.
XX
PA  (MAXY-) MAXYGEN APS.
XX
PI  Hansen CK;
XX
PT  Dispersions of particles for administration to the lung and circulatory
PT  system by inhalation, e.g. from a nebulizer.
XX
PS  Claim 22; Page 138-139; 139pp; English.
XX
CC  The present invention relates to novel dispersions of a conjugate
```

CC   comprising at least one non-polypeptide moiety covalently attached to a
CC   polypeptide variant, wherein the amino acid sequence of the polypeptide
CC   variant differs from that of the corresponding wild-type human
CC   polypeptide in that at least 1 residue comprising an attachment group for
CC   the non-polypeptide group has been introduced and/or removed. The wild-
CC   type human polypeptide is selected from the group consisting of
CC   interferon-gamma (IFNG), interferon-beta, interferon alpha, follicle
CC   stimulating hormone, (FSH), interleukin 10 (IL-10) and granulocyte colony
CC   stimulating factor (G-CSF). The dispersion is used for the preparation of
CC   a medicament for pulmonary and/or circulatory administration. The present
CC   sequence is human wild type granulocyte colony stimulating factor (G-CSF)
XX
SQ   Sequence 174 AA;

Query Match   100.0%;   Score 896;   DB 4;   Length 174;
Best Local Similarity   100.0%;   Pred. No. 3.8e-90;
Matches 174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 15
AAU79877
ID   AAU79877 standard; protein; 174 AA.
XX
AC   AAU79877;
XX
DT   15-JUL-2002 (first entry)
XX
DE   Human granulocyte colony-stimulating factor (G-CSF).
XX
KW   Human; granulocyte colony-stimulating factor; G-CSF; vasotropic;
KW   vulnerary; immunosuppressive; cerebroprotective; nephrotropic;
KW   respiratory; cardiant; antiarteriosclerotic; ischaemic disease;
KW   external wound; transplant rejection; ischaemic cerebrovascular disorder;
KW   cerebral apoplexy; cerebral infarction; ischaemic kidney disease;
KW   ischaemic pulmonary disease; coronary occlusion; ischaemic limb disease;
KW   ischaemic heart disease; myocardial ischaemia; myocardial infarction;
KW   cardiac insufficiency; arteriosclerosis; kinesitherapy; pharmacotherapy;
KW   haemokinesis reconstruction; arteriosclerosis obliteran; gene therapy;
KW   intramuscular transplantation.
XX
OS   Homo sapiens.
XX
PN   WO200222163-A1.
XX
PD   21-MAR-2002.
XX
PF   13-SEP-2001; 2001WO-JP007946.
XX
PR   13-SEP-2000; 2000JP-00277562.
XX
PA   (CHUS) CHUGAI SEIYAKU KK.
XX
PI   Miyai T, Tamura M;
XX
DR   WPI; 2002-371949/40.
XX
PT   Remedies for ischemic diseases comprise human granulocyte colony-
PT   stimulating factor.
XX
PS   Disclosure; Page 4-5; 21pp; Japanese.
XX

CC   The invention describes remedies comprising a human granulocyte colony-
CC   stimulating factor (G-CSF). Used for treating and preventing ischaemic
CC   diseases such as external wounds, transplant rejection, ischaemic
CC   cerebrovascular disorders (e.g. cerebral apoplexy and cerebral
CC   infarction), ischaemic kidney diseases, ischaemic pulmonary diseases,
CC   ischaemia due to coronary occlusion, ischaemic limb diseases, ischaemic
CC   heart disease (e.g. myocardial ischaemia, myocardial infarction or
CC   cardiac insufficiency) or arteriosclerosis. Agents may also be used for
CC   overcoming problems occurring particularly in the conventional
CC   kinesitherapy, pharmacotherapy and haemokinesis reconstruction for
CC   arteriosclerosis obliterans, gene therapy and intramuscular
CC   transplantation. This is the amino acid sequence of the human
CC   granulocytic colony-stimulating factor (G-CSF)
XX
SQ   Sequence 174 AA;

Query Match   100.0%;   Score 896;   DB 5;   Length 174;
Best Local Similarity   100.0%;   Pred. No. 3.8e-90;
Matches 174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

Search completed: October 6, 2004, 16:04:03
Job time : 129 secs

```
              GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:   October  6, 2004, 16:02:01 ; Search time 32 Seconds
                                       (without alignments)
                                       280.716 Million cell updates/sec

Title:       US-10-009-792C-19
Perfect score:  896
Sequence:    1 TPLGPASSLPQSFLKCLEQ.........SHLQSFLEVSYRVLRHLAQP 174

Scoring table:  BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:   389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:        389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
             1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
             2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
             3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
             4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
             5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
             6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	896	100.0	174	1	US-08-225-224-5	Sequence 5, Appli
2	896	100.0	174	1	US-08-431-459A-31	Sequence 31, Appl
3	896	100.0	174	3	US-08-722-258-5	Sequence 5, Appli
4	896	100.0	174	3	US-09-221-181-1	Sequence 1, Appli
5	896	100.0	174	4	US-09-904-196B-1	Sequence 1, Appli
6	896	100.0	174	4	US-09-462-941-6	Sequence 6, Appli
7	896	100.0	174	4	US-09-479-313B-2	Sequence 2, Appli
8	896	100.0	174	4	US-09-760-008A-1	Sequence 1, Appli
9	896	100.0	174	5	PCT-US95-04468-5	Sequence 5, Appli
10	896	100.0	174	6	5194592-26	Patent No. 5194592
11	896	100.0	175	1	US-08-010-099-2	Sequence 2, Appli
12	896	100.0	175	1	US-08-167-721-1	Sequence 1, Appli
13	896	100.0	175	1	US-08-428-732-6	Sequence 6, Appli
14	896	100.0	175	2	US-08-448-716-2	Sequence 2, Appli
15	896	100.0	175	2	US-08-321-510-2	Sequence 2, Appli
16	896	100.0	175	2	US-08-570-943-1	Sequence 1, Appli
17	896	100.0	175	2	US-08-879-760-2	Sequence 2, Appli
18	896	100.0	175	3	US-08-505-187-4	Sequence 4, Appli
19	896	100.0	175	3	US-09-304-186-2	Sequence 2, Appli
20	896	100.0	175	4	US-09-230-733-1	Sequence 1, Appli
21	896	-100.0	175	4	US-09-479-313B-15	Sequence 15, Appl
22	896	100.0	175	4	US-09-754-532-2	Sequence 2, Appli
23	896	100.0	175	5	PCT-US95-01729-2	Sequence 2, Appli
24	896	100.0	176	3	US-08-469-318-161	Sequence 161, App
25	896	100.0	176	3	US-08-468-609A-161	Sequence 161, App
26	896	100.0	176	4	US-08-446-872A-161	Sequence 161, App
27	896	100.0	176	4	US-08-762-227A-161	Sequence 161, App
28	896	100.0	176	5	PCT-US95-01185-161	Sequence 161, App
29	896	100.0	177	2	US-08-797-689-14	Sequence 14, Appl
30	896	100.0	177	4	US-09-984-186-14	Sequence 14, Appl
31	896	100.0	204	1	US-09-792-019B-10	Sequence 10, Appl
32	896	100.0	204	3	US-08-988-819-10	Sequence 10, Appl
33	896	100.0	204	3	US-09-016-534-10	Sequence 10, Appl
34	896	100.0	204	3	US-09-097-869-5	Sequence 5, Appli
35	896	100.0	307	3	US-08-469-318-121	Sequence 121, App
36	896	100.0	307	3	US-08-469-318-134	Sequence 134, App
37	896	100.0	307	3	US-08-469-318-146	Sequence 146, App
38	896	100.0	307	3	US-08-469-318-147	Sequence 147, App
39	896	100.0	307	3	US-08-468-609A-121	Sequence 121, App
40	896	100.0	307	3	US-08-468-609A-134	Sequence 134, App
41	896	100.0	307	3	US-08-468-609A-146	Sequence 146, App
42	896	100.0	307	3	US-08-468-609A-147	Sequence 147, App
43	896	100.0	307	4	US-08-446-872A-121	Sequence 121, App
44	896	100.0	307	4	US-08-446-872A-134	Sequence 134, App
45	896	100.0	307	4	US-08-446-872A-146	Sequence 146, App

ALIGNMENTS

```
RESULT 1
US-08-225-224-5
; Sequence 5, Application US/08225224
; Patent No. 5635599
;  GENERAL INFORMATION:
;    APPLICANT: PASTAN, Ira
;    APPLICANT: KREITMAN, Robert J.
;    TITLE OF INVENTION:  CIRCULARLY PERMUTATED LIGANDS AND
;    TITLE OF INVENTION:  CIRCULARLY PERMUTED FUSION PROTEINS
;    NUMBER OF SEQUENCES: 57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Townsend and Townsend Khourie and Crew
;      STREET: Steuart Street Tower, One Market Plaza
;      CITY: San Francisco
;      STATE: California
;      COUNTRY: US
;      ZIP: 94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/225,224
;      FILING DATE: 8-APR-1994
;      CLASSIFICATION: 530
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Weber, Ellen L.
;      REGISTRATION NUMBER: 32,762
;      REFERENCE/DOCKET NUMBER: 15280-193
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:    (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 174 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: unknown
;      TOPOLOGY: unknown
;    MOLECULE TYPE: protein
;    FEATURE:
;      NAME/KEY: Protein
;      LOCATION: 1..174
;      OTHER INFORMATION: /label= G-CSF
US-08-225-224-5

Query Match             100.0%;  Score 896;  DB 1;  Length 174;
Best Local Similarity   100.0%;  Pred. No. 4.3e-92;
Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 2
US-08-431-459A-31
; Sequence 31, Application US/08431459A
; Patent No.5840543
; GENERAL INFORMATION:
; APPLICANT: Hockney, Robert C.
; APPLICANT: Kara, Bhuphendra V.
; TITLE OF INVENTION: FERMENTATION PROCESS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44M diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,459A
; FILING DATE: 01-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,676
; FILING DATE: 22-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/661,306
; FILING DATE: 27-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9004390.2
; FILING DATE: 27-FEB-1990
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-431-459A-31

Query Match   100.0%; Score 896; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 4.3e-92;
Matches 174; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 3
US-08-722-258-5

; Sequence 5, Application US/08722258
; Patent No. 6011002
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Kreitman, Robert J.
; APPLICANT: Puri, Raj K.
; TITLE OF INVENTION: Circularly Permuted Ligands and
; TITLE OF INVENTION: Circularly Permuted Chimeric Molecules
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,258
; FILING DATE: 08-JAN-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US95/04468
; FILING DATE: 06-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,224
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-19310OUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..174
; OTHER INFORMATION: /note= "granulocyte
; OTHER INFORMATION: colony-stimulating factor (G-CSF)"
US-08-722-258-5

Query Match   100.0%; Score 896; DB 3; Length 174;
Best Local Similarity 100.0%; Pred. No. 4.3e-92;
Matches 174; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 4
US-09-221-181-1
; Sequence 1, Application US/09221181
; Patent No. 6245740

; GENERAL INFORMATION:
; APPLICANT: GOLDENBERG, MERRILL
; APPLICANT: SHAN, DAXIAN
; APPLICANT: BEEKMAN, ALICE
; TITLE OF INVENTION: POLYOL/OIL SUSPENSIONS FOR SUSTAINED RELEASE OF
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: A-576
; CURRENT APPLICATION NUMBER: US/09/221,181
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 174
; TYPE: PRT
; ORGANISM: granulocyte colony-stimulating factor
US-09-221-181-1

Query Match   100.0%; Score 896; DB 3; Length 174;
Best Local Similarity   100.0%; Pred. No. 4.3e-92;
Matches 174; Conservative   0; Mismatches   0;   Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 5
US-09-904-196B-1
; Sequence 1, Application US/09904196B
; Patent No. 6555660
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF CONJUGATES
; FILE REFERENCE: 31-000700US
; CURRENT APPLICATION NUMBER: US/09/904,196B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US/09/760,008
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-904-196B-1

Query Match   100.0%; Score 896; DB 4; Length 174;
Best Local Similarity   100.0%; Pred. No. 4.3e-92;
Matches 174; Conservative   0; Mismatches   0;   Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 6
US-09-462-941-6
; Sequence 6, Application US/09462941
; Patent No. 6608183
; GENERAL INFORMATION:
; APPLICANT: Cox III, George N
; APPLICANT: Bolder Biotechnology, Inc.
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4152-1-PUS
; CURRENT APPLICATION NUMBER: US/09/462,941
; CURRENT FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-462-941-6

Query Match   100.0%; Score 896; DB 4; Length 174;
Best Local Similarity   100.0%; Pred. No. 4.3e-92;
Matches 174; Conservative   0; Mismatches   0;   Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 7
US-09-479-313B-2
; Sequence 2, Application US/09479313B
; Patent No. 6627186
; GENERAL INFORMATION:
; APPLICANT: Dahiyat, Bassil I.
; APPLICANT: Luo, Peizhi
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEIN VARIANTS OF hG-CSF WITH GRANULOPOIETIC
; TITLE OF INVENTION: ACTIVITY
; FILE REFERENCE: A-67614-2/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/479,313B
; CURRENT FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: US 60/115,131
; PRIOR FILING DATE: 1999-01-06
; PRIOR APPLICATION NUMBER: US 60/118,831
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-479-313B-2

Query Match   100.0%; Score 896; DB 4; Length 174;
Best Local Similarity   100.0%; Pred. No. 4.3e-92;
Matches 174; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 8
US-09-760-008A-1
; Sequence 1, Application US/09760008A
; Patent No. 6646110
; GENERAL INFORMATION:
;   APPLICANT: NISSEN, TORBEN LAUSGAARD
;   APPLICANT: ANDERSEN, KIM VILBOUR
;   APPLICANT: HANSEN, CHRISTIAN KARSTEN
;   APPLICANT: MIKKELSEN, JAN MOLLER
;   TITLE OF INVENTION: G-CSF CONJUGATES.
;   FILE REFERENCE: 31-000700US
;   CURRENT APPLICATION NUMBER: US/09/760,008A
;   CURRENT FILING DATE: 2001-05-11
;   PRIOR APPLICATION NUMBER: 60/176,376
;   PRIOR APPLICATION NUMBER: 60/189,506
;   PRIOR FILING DATE: 2000-01-14
;   PRIOR FILING DATE: 2000-03-15
;   PRIOR APPLICATION NUMBER: 60/215,644
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DK PA 2000 00024
;   PRIOR FILING DATE: 2000-01-10
;   PRIOR APPLICATION NUMBER: DK PA 2000 00341
;   PRIOR FILING DATE: 2000-03-02
;   PRIOR APPLICATION NUMBER: DK PA 2000 00943
;   PRIOR FILING DATE: 2000-06-16
;   NUMBER OF SEQ ID NOS: 15
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 174
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-760-008A-1

Query Match   100.0%; Score 896; DB 4; Length 174;
Best Local Similarity   100.0%; Pred. No. 4.3e-92;
Matches 174; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 9
PCT-US95-04468-5
; Sequence 5, Application PC/TUS9504468
; GENERAL INFORMATION:
;   APPLICANT:
;   TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND
;   TITLE OF INVENTION: CIRCULARLY PERMUTED FUSION PROTEINS
;   NUMBER OF SEQUENCES: 59
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US95/04468
;   FILING DATE: 07-APR-1995
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/225,224
;   FILING DATE: 08-APR-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Weber, Ellen L.
;   REGISTRATION NUMBER: 32,762
;   REFERENCE/DOCKET NUMBER: 15280-193-1PC
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 543-9600
;   TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 174 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: protein
;   FEATURE:
;   NAME/KEY: Protein
;   LOCATION: 1..174
;   OTHER INFORMATION: /label= G-CSF
PCT-US95-04468-5

Query Match   100.0%; Score 896; DB 5; Length 174;
Best Local Similarity   100.0%; Pred. No. 4.3e-92;
Matches 174; Conservative   0; Mismatches   0; Indels   .0; Gaps   0;

```
Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 10
5194592-26
;Patent No. 5194592
;   APPLICANT: YOSHIDA, HAJIME
;   TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO NOVEL
;POLYPEPTIDES DERIVTIVES OF HUMAN GRANULOCYTE COLONY
;STIMULATING FACTOR
;   NUMBER OF SEQUENCES: 83
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/318,527
;   FILING DATE: 3-MAR-1989
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 136,647
;   FILING DATE: 22-DEC-1987
;SEQ ID NO:26:
;   LENGTH: 174
5194592-26

Query Match   100.0%; Score 896; DB 6; Length 174;
Best Local Similarity   100.0%; Pred. No. 4.3e-92;

Matches 174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 11
US-08-010-099-2
; Sequence 2, Application US/080:0099
; Patent No. 5581476
GENERAL INFORMATION:
;   APPLICANT: Osslund, Timothy
;   TITLE OF INVENTION: G-CSF ANALOG COMPOSITIONS AND METHODS
;   NUMBER OF SEQUENCES: 110
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Amgen, Inc.
;   STREET: Amgen Center, 1840 DeHavilland Drive
;   CITY: Thousand Oaks
;   STATE: California
;   COUNTRY: USA
;   ZIP: 91320-1789
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/010,099
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Pessin, Karol
;   REGISTRATION NUMBER: 34,899
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 805/499-5725
;   TELEFAX: 805/499-8011
;   INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 175 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-010-099-2

Query Match   100.0%;   Score 896;   DB 1;   Length 175;
Best Local Similarity   100.0%;   Pred. No. 4.4e-92;
Matches 174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 61

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 121

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 175
```

RESULT 12
US-08-167-721-1
; Sequence 1, Application US/08167721

; Patent No. 5597562
GENERAL INFORMATION:
;   APPLICANT: No. 5597562ura, Hideaki
;   APPLICANT: Maruyama, Kazutoshi
;   TITLE OF INVENTION: Oral Dosage Form of Biologically Active
;   TITLE OF INVENTION: Proteins
;   NUMBER OF SEQUENCES: 1
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;   ADDRESSEE: Bicknell
;   STREET: Two First National Plaza, 20 South Clark
;   STREET: Street
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60603
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/167,721
;   FILING DATE:
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/07/994,076
;   FILING DATE:
;   APPLICATION NUMBER: US/07/709,622
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Borun, Michael F.
;   REGISTRATION NUMBER: 25447
;   REFERENCE/DOCKET NUMBER: 11009/30403
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (312) 346-5750
;   TELEFAX: (312) 984-9740
;   TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 175 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-167-721-1

Query Match   100.0%;   Score 896;   DB 1;   Length 175;
Best Local Similarity   100.0%;   Pred. No. 4.4e-92;
Matches 174;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 61

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 121

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 175
```

RESULT 13
US-08-428-732-6
; Sequence 6, Application US/08428732
; Patent No. 5606024
GENERAL INFORMATION:
;   APPLICANT: Boone, Thomas C.
;   APPLICANT: Miller, Allan L.
;   TITLE OF INVENTION: DNA Encoding Canine Granulocyte Colony
;   TITLE OF INVENTION: Stimulating Factor (G-CSF)
;   NUMBER OF SEQUENCES: 6

```
                  CORRESPONDENCE ADDRESS:
;                   ADDRESSEE:  Amgen Inc.
;                   STREET:  1840 Dehavilland Drive
;                   CITY:  Thousand Oaks
;                   STATE:  California
;                   COUNTRY:  USA
;                   ZIP:  91320-1789
;                 COMPUTER READABLE FORM:
;                   MEDIUM TYPE:  Floppy disk
;                   COMPUTER:  IBM PC compatible
;                   OPERATING SYSTEM:  PC-DOS/MS-DOS
;                   SOFTWARE:  PatentIn Release #1.0,  Version #1.25
;                 CURRENT APPLICATION DATA:
;                   APPLICATION NUMBER:  US/08/428,732
;                   FILING DATE:
;                   CLASSIFICATION:
;                 ATTORNEY/AGENT INFORMATION:
;                   NAME: Crandall, Craig A.
;                   REFERENCE/DOCKET NUMBER:  A-173-C2
;               INFORMATION FOR SEQ ID NO:  6:
;                 SEQUENCE CHARACTERISTICS:
;                   LENGTH:  175 amino acids
;                   TYPE:  amino acid
;                   STRANDEDNESS:  single
;                   TOPOLOGY:  linear
;                   MOLECULE TYPE:  protein
US-08-428-732-6

Query Match             100.0%;    Score 896;  DB 1;  Length 175;
Best Local Similarity   100.0%;    Pred. No. 4.4e-92;
Matches  174;   Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 61

Qy   61  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 121

Qy  121  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 175

RESULT 14
US-08-448-716-2
; Sequence 2, Application US/08448716
; Patent No. 5790421
; GENERAL INFORMATION:
;   APPLICANT:  Osslund, Timothy D.
;   TITLE OF INVENTION:  G-CSF ANALOG COMPOSITIONS AND METHODS
;   NUMBER OF SEQUENCES:  110
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Amgen Inc.
;     STREET:  Amgen Center, 1840 DeHavilland Drive
;     CITY:  Thousand Oaks
;     STATE:  California
;     COUNTRY:  United States of America
;     ZIP:  91320-1789
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0,  Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/448,716
;     FILING DATE:  24-MAY-1995
;     CLASSIFICATION:  530
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Pessin, Karol
;     REGISTRATION NUMBER:  34,899
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  805/499-5725
;     TELEFAX:  805/499-8011
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  175 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
US-08-448-716-2

Query Match             100.0%;    Score 896;  DB 1;  Length 175;
Best Local Similarity   100.0%;    Pred. No. 4.4e-92;
Matches  174;   Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 61

Qy   61  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 121

Qy  121  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 175

RESULT 15
US-08-321-510-2
; Sequence 2, Application US/08321510
; Patent No. 5824784
; GENERAL INFORMATION:
;   APPLICANT:  Kinstler, Olaf B.
;   APPLICANT:  Gabriel, Nancy E.
;   APPLICANT:  Farrar, Christine E.
;   APPLICANT:  DePrince, Randolph B.
;   TITLE OF INVENTION:  N-Terminally Chemically Modified Protein
;   TITLE OF INVENTION:  Composition and Methods
;   NUMBER OF SEQUENCES:  2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Amgen Inc.
;     STREET:  1840 Dehavilland Drive
;     CITY:  Thousand Oaks
;     STATE:  California
;     COUNTRY:  USA
;     ZIP:  91320
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0,  Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/321,510
;     FILING DATE:
;     CLASSIFICATION:  424
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Pessin, Karol M.
;     REFERENCE/DOCKET NUMBER:  A-286
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  175 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
US-08-321-510-2

Query Match             100.0%;    Score 896;  DB 2;  Length 175;
Best Local Similarity   100.0%;    Pred. No. 4.4e-92;
Matches  174;   Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP   61

Qy      61  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ  121

Qy     121  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP  174
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP  175
```

Search completed: October 6, 2004, 16:07:55
Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:06:41 ;Search time 126 Seconds
                                        (without alignments)
                                        444.389 Million cell updates/sec

Title: US-10-009-792C-19
Perfect score: 896
Sequence: 1 TPLGPASSLPQSFLKCLEQ.........SHLQSFLEVSYRVLRHLAQP 174

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched: 1351062 segs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   Published_Applications_AA:*
    1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
    2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
    3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
    4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
    5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
    6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
    7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
    8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
    9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
   10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
   11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
   12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
   13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
   14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
   15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
   16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
   17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
   18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	896	100.0	174	9	US-09-760-008A-1	Sequence 1, Appli
2	896	100.0	174	9	US-09-950-473-2	Sequence 2, Appli
3	896	100.0	174	9	US-09-921-114-1	Sequence 2, Appli
4	896	100.0	174	10	US-09-950-123-2	Sequence 2, Appli
5	896	100.0	174	12	US-10-411-037-2	Sequence 2, Appli
6	896	100.0	174	12	US-10-411-026-2	Sequence 2, Appli
7	896	100.0	174	12	US-10-467-396-1	Sequence 1, Appli
8	896	100.0	174	12	US-10-436-784-1	Sequence 1, Appli
9	896	100.0	174	13	US-10-016-403-1	Sequence 1, Appli
10	896	100.0	174	13	US-10-003-496-1	Sequence 1, Appli
11	896	100.0	174	14	US-10-192-294-1	Sequence 19, Appli
12	896	100.0	174	14	US-10-009-792A-19	Sequence 1, Appli
13	896	100.0	174	14	US-10-318-966-1	Sequence 1, Appli
14	896	100.0	174	14	US-10-400-377-6	Sequence 6, Appli
15	896	100.0	174	14	US-10-400-708-6	Sequence 6, Appli
16	896	100.0	174	14	US-10-298-148-6	Sequence 6, Appli
17	896	100.0	174	16	US-10-410-962-2	Sequence 2, Appli
18	896	100.0	174	16	US-10-411-049-2	Sequence 2, Appli
19	896	100.0	174	16	US-10-632-695-2	Sequence 2, Appli
20	896	100.0	174	16	US-10-410-930-2	Sequence 2, Appli
21	896	100.0	174	16	US-10-410-997-2	Sequence 2, Appli
22	896	100.0	174	16	US-10-411-012-2	Sequence 2, Appli
23	896	100.0	174	16	US-10-287-994-2	Sequence 2, Appli
24	896	100.0	174	16	US-10-410-913-2	Sequence 2, Appli
25	896	100.0	174	16	US-10-659-295-37	Sequence 37, Appl
26	896	100.0	174	16	US-10-750-797-1	Sequence 1, Appli
27	896	100.0	174	16	US-10-773-939-6	Sequence 6, Appli
28	896	100.0	174	16	US-10-774-149-6	Sequence 6, Appli
29	896	100.0	174	16	US-10-468-496-205	Sequence 205, App
30	896	100.0	175	9	US-09-754-532-2	Sequence 1, Appli
31	896	100.0	175	9	US-09-230-733-1	Sequence 1, Appli
32	896	100.0	175	9	US-09-921-114-2	Sequence 2, Appli
33	896	100.0	175	12	US-10-436-784-2	Sequence 2, Appli
34	896	100.0	175	12	US-09-817-725-2	Sequence 2, Appli
35	896	100.0	175	14	US-10-131-956-2	Sequence 2, Appli
36	896	100.0	175	14	US-10-264-846-2	Sequence 2, Appli
37	896	100.0	175	14	US-10-009-792A-21	Sequence 21, Appl
38	896	100.0	175	14	US-10-032-108-2	Sequence 2, Appli
39	896	100.0	175	14	US-10-345-639A-2	Sequence 2, Appli
40	896	100.0	175	14	US-10-365-418-1	Sequence 1, Appli
41	896	100.0	175	15	US-10-168-956A-3	Sequence 3, Appli
42	896	100.0	175	15	US-10-632-695-15	Sequence 15, Appl
43	896	100.0	175	16	US-10-659-295-38	Sequence 38, Appl
44	896	100.0	175	16	US-10-750-797-2	Sequence 2, Appli
45	896	100.0	176	14	US-10-083-446-161	Sequence 161, App

ALIGNMENTS

RESULT 1
US-09-760-008A-1
; Sequence 1, Application US/09760008A
; Patent No. US2002004483A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF CONJUGATES
; FILE REFERENCE: 31-000700US
; CURRENT APPLICATION NUMBER: US/09/760,008A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-008A-1

Query Match   100.0%; Score 896; DB 9; Length 174;
Best Local Similarity   100.0%; Pred. No. 2.3e-85;
Matches 174; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

```
Db        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP  60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFIYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFIYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 2
US-09-950-473-2
; Sequence 2, Application US/09950473
; Patent No. US2002015148BA1
; GENERAL INFORMATION:
; APPLICANT: Sarkar, Casim
; APPLICANT: Lauffenburger, Douglas
; TITLE OF INVENTION: G-CSF Analog Compositions and Methods
; FILE REFERENCE: 01017/37732
; CURRENT APPLICATION NUMBER: US/09/950,473
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-473-2

Query Match   100.0%;   Score 896; DB 9; Length 174;
Best Local Similarity 100.0%; Pred. No. 2.3e-85;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP  60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFIYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFIYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 3
US-09-921-114-1
; Sequence 1, Application US/09921114
; Patent No. US2002017768BA1
; GENERAL INFORMATION:
; APPLICANT: ISHIKAWA ET AL
; TITLE OF INVENTION: CHEMICALLY-MODIFIED G-CSF
; FILE REFERENCE: 11009/36193A
; CURRENT APPLICATION NUMBER: US/09/921,114
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 09/518,896
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 08/957,719
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: US 07/983,620
; PRIOR FILING DATE: 1992-11-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-114-1

Query Match   100.0%;   Score 896; DB 9; Length 174;

Best Local Similarity 100.0%; Pred. No. 2.3e-85;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP  60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFIYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFIYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 4
US-09-950-123-2
; Sequence 2, Application US/09950123
; Publication No. US2003016652BA1
; GENERAL INFORMATION:
; APPLICANT: Sarkar, Casim
; APPLICANT: Lauffenburger, Douglas
; APPLICANT: Tidor, Bruce
; TITLE OF INVENTION: G-CSF Analog Compositions and Methods
; FILE REFERENCE: 01017/37377
; CURRENT APPLICATION NUMBER: US/09/950,123
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-123-2

Query Match   100.0%;   Score 896; DB 10; Length 174;
Best Local Similarity 100.0%; Pred. No. 2.3e-85;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP  60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFIYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFIYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 5
US-10-411-037-2
; Sequence 2, Application US/10411037
; Publication No. US2004043446A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryn
; TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPHA
; TITLE OF INVENTION: GALACTOSIDASE A
; FILE REFERENCE: 040853-01-5082
; CURRENT APPLICATION NUMBER: US/10/411,037
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692

```
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-037-2

Query Match             100.0%;  Score 896;  DB 12;  Length 174;
Best Local Similarity   100.0%;  Pred. No. 2.3e-85;
Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 6
US-10-411-026-2
; Sequence 2, Application US/10411026
; Publication No. US20040063911A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: PROTEIN REMODELING METHODS AND PROTEINS/PEPTIDES PRODUCED BY THE
; TITLE OF INVENTION: METHODS
; FILE REFERENCE: 040853-01-5053
; CURRENT APPLICATION NUMBER: US/10/411,026
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-026-2


Query Match             100.0%;  Score 896;  DB 12;  Length 174;
Best Local Similarity   100.0%;  Pred. No. 2.3e-85;
Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 7
US-10-467-396-1
; Sequence 1, Application US/10467396
; Publication No. US20040062749A1
; GENERAL INFORMATION:
; APPLICANT: Carr, Francis J.
; APPLICANT: Carter, Graham
; APPLICANT: Jones, Tim
; APPLICANT: Williams, Stephen
; TITLE OF INVENTION: MODIFIED GRANULOCYTE COLONY STIMULATING
; TITLE OF INVENTION: FACTOR (G-CSF) WITH REDUCED IMMUNOGENICITY
; FILE REFERENCE: MER-113
; CURRENT APPLICATION NUMBER: US/10/467,396
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: EP 01102617.6
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: EP 01103954.2
; PRIOR FILING DATE: 2001-02-19
; PRIOR APPLICATION NUMBER: PCT/EP/02/01171
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-467-396-1


Query Match             100.0%;  Score 896;  DB 12;  Length 174;
Best Local Similarity   100.0%;  Pred. No. 2.3e-85;
Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 8
US-10-436-784-1
; Sequence 1, Application US/10436784
; Publication No. US20030204057A1
; GENERAL INFORMATION:
; APPLICANT: ISHIKAWA ET AL
; TITLE OF INVENTION: CHEMICALLY-MODIFIED G-CSF
; FILE REFERENCE: 11009/36193A
; CURRENT APPLICATION NUMBER: US/10/436,784
; CURRENT FILING DATE: 2003-05-12
```

```
;       PRIOR APPLICATION NUMBER: US 09/518,896
;       PRIOR APPLICATION FILING DATE: 2000-03-06
;       PRIOR APPLICATION NUMBER: US 08/957,719
;       PRIOR APPLICATION FILING DATE: 1997-10-27
;       PRIOR APPLICATION NUMBER: US 07/983,620
;       PRIOR APPLICATION FILING DATE: 1992-11-30
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
;
;  SEQ ID NO 1
;    LENGTH: 174
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-436-784-1

Query Match             100.0%;    Score 896;  DB 12;  Length 174;
Best Local Similarity   100.0%;    Pred. No. 2.3e-85;
  Matches 174;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy        61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy       121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 9
US-10-016-403-1
;  Sequence 1, Application US/10016403
;  Publication No. US20020107505A1
;  GENERAL INFORMATION:
;     APPLICANT: Holladay, Leslie A.
;     TITLE OF INVENTION: MODIFICATION OF POLYPEPTIDE DRUGS TO
;        INCREASE ELECTROTRANSPORT FLUX
;
;  NUMBER OF SEQUENCES: 10
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Stroud, Stroud, Willink, Thompson & Howard
;     STREET: 25 West Main Street
;     CITY: Madison
;     STATE: WI
;     COUNTRY: USA
;     ZIP: 53701-2236
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/10/016,403
;     FILING DATE: 10-Dec-2001
;     CLASSIFICATION: <Unknown>
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/466,610
;     FILING DATE: 1995-JUN-06
;  ATTORNEY/AGENT INFORMATION:
;     NAME: Frenchick, Grady J.
;     REGISTRATION NUMBER: 29,018
;     REFERENCE/DOCKET NUMBER: 8734.28
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 608-257-2281
;     TELEFAX: 608-257-7643
;  INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 174 amino acids
;        TYPE: amino acid
;        TOPOLOGY: linear
;     FEATURE:
;        NAME/KEY: Peptide
```

```
;        LOCATION: 1..174
;        OTHER INFORMATION: /note= "granulocyte-colony
;           stimulating factor"
;  SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-016-403-1

Query Match             100.0%;    Score 896;  DB 13;  Length 174;
Best Local Similarity   100.0%;    Pred. No. 2.3e-85;
  Matches 174;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy        61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy       121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 10
US-10-003-496-1
;  Sequence 1, Application US/10003496
;  Publication No. US20020142964A1
;  GENERAL INFORMATION:
;     APPLICANT: Maxygen ApS
;     APPLICANT: Maxygen Holdings Ltd.
;     TITLE OF INVENTION: Single-Chain Polypeptides
;     FILE REFERENCE: 0218us210
;     CURRENT APPLICATION NUMBER: US/10/003,496
;     CURRENT FILING DATE: 2002-01-31
;     PRIOR APPLICATION NUMBER: US 60/245,727
;     PRIOR FILING DATE: 2000-11-02
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
;
;  SEQ ID NO 1
;    LENGTH: 174
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-003-496-1

Query Match             100.0%;    Score 896;  DB 13;  Length 174;
Best Local Similarity   100.0%;    Pred. No. 2.3e-85;
  Matches 174;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy        61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy       121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174

RESULT 11
US-10-192-294-1
;  Sequence 1, Application US/10192294
;  Publication No. US20030118612A1
;  GENERAL INFORMATION:
;     APPLICANT: NISSEN, Torben Lauesgaard
;     TITLE OF INVENTION: G-CSF Conjugates
;     FILE REFERENCE: 0258us310
;     CURRENT APPLICATION NUMBER: US/10/192,294
;     CURRENT FILING DATE: 2002-07-10
;     PRIOR APPLICATION NUMBER: US 09/904,196
;     PRIOR FILING DATE: 2001-07-11
```

```
;   PRIOR APPLICATION NUMBER: DK PA 2002 00447
;   PRIOR FILING DATE: 2002-03-22
;   PRIOR APPLICATION NUMBER: DK PA 2002 00708
;   PRIOR FILING DATE: 2002-05-08
;   NUMBER OF SEQ ID NOS: 15
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 1
;   LENGTH: 174
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-192-294-1

  Query Match         100.0%;  Score 896;  DB 14;  Length 174;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-85;
  Matches  174;  Conservative    0;  Mismatches   0;  Indels     0;  Gaps    0;

Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174


RESULT 12
US-10-009-792A-19
; Sequence 19, Application US/10009792A
; Publication No. US20030153049A1
; GENERAL INFORMATION:
;   APPLICANT: LEE, Sang-Yup
;   APPLICANT: JEONG, Ki-Jun
;   TITLE OF INVENTION: ESCHERICHIA COLI STRAIN SECRETING HUMAN
;   TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR (G-CSF)
;   FILE REFERENCE: HYLEE60.001APC
;   CURRENT APPLICATION NUMBER: US/10/009,792A
;   CURRENT FILING DATE: 2002-10-29
;   PRIOR APPLICATION NUMBER: PCT/KR01/00549
;   PRIOR FILING DATE: 2001-03-31
;   PRIOR APPLICATION NUMBER: KR 10-2000-0017052
;   PRIOR FILING DATE: 2000-03-31
;   NUMBER OF SEQ ID NOS: 27
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 19
;   LENGTH: 174
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-009-792A-19

  Query Match         100.0%;  Score 896;  DB 14;  Length 174;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-85;
  Matches  174;  Conservative    0;  Mismatches   0;  Indels     0;  Gaps    0;

Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174


RESULT 13
US-10-318-966-1
; Sequence 1, Application US/10318966
```

```
; Publication No. US20030158375A1
; GENERAL INFORMATION:
;   APPLICANT: NISSEN, TORBEN LAUESGAARD
;   APPLICANT: ANDERSEN, KIM VILBOUR
;   APPLICANT: HANSEN, CHRISTIAN KARSTEN
;   APPLICANT: MIKKELSEN, JAN MOLLER
;   APPLICANT: SCHAMBYE, HANS THALSGARD
;   TITLE OF INVENTION: G-CSF POLYPEPTIDES AND CONJUGATES
;   FILE REFERENCE: 0208us620
;   CURRENT APPLICATION NUMBER: US/10/318,966
;   CURRENT FILING DATE: 2002-12-13
;   PRIOR APPLICATION NUMBER: 09/904,196
;   PRIOR FILING DATE: 2001-07-11
;   PRIOR APPLICATION NUMBER: 09/760,008
;   PRIOR FILING DATE: 2001-01-10
;   PRIOR APPLICATION NUMBER: 60/176,376
;   PRIOR FILING DATE: 2000-01-14
;   PRIOR APPLICATION NUMBER: 60/189,506
;   PRIOR FILING DATE: 2000-03-15
;   PRIOR APPLICATION NUMBER: 60/215,644
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DK PA 2000 00024
;   PRIOR FILING DATE: 2000-01-10
;   PRIOR APPLICATION NUMBER: DK PA 2000 00341
;   PRIOR FILING DATE: 2000-03-02
;   PRIOR APPLICATION NUMBER: DK PA 2000 00943
;   PRIOR FILING DATE: 2000-06-16
;   NUMBER OF SEQ ID NOS: 15
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 1
;   LENGTH: 174
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-318-966-1

  Query Match         100.0%;  Score 896;  DB 14;  Length 174;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-85;
  Matches  174;  Conservative    0;  Mismatches   0;  Indels     0;  Gaps    0;

Qy     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174


RESULT 14
US-10-400-377-6
; Sequence 6, Application US/10400377
; Publication No. US20030162249A1
; GENERAL INFORMATION:
;   APPLICANT: Cox III, George N
;   APPLICANT: Bolder Biotechnology, Inc.
;   TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
;   FILE REFERENCE: 4152-1-PUS
;   CURRENT APPLICATION NUMBER: US/10/400,377
;   CURRENT FILING DATE: 2003-03-26
;   PRIOR APPLICATION NUMBER: US/09/462,941
;   PRIOR FILING DATE: 2000-01-14
;   PRIOR APPLICATION NUMBER: 60/052,516
;   PRIOR FILING DATE: 1997-07-14
;   NUMBER OF SEQ ID NOS: 41
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 6
;   LENGTH: 174
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

US-10-400-377-6

```
Query Match        100.0%;   Score 896;  DB 14;   Length 174;
Best Local Similarity 100.0%;   Pred. No. 2.3e-85;
Matches 174; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

RESULT 15
US-10-400-708-6
; Sequence 6, Application US/10400708
; Publication No. US20030166865A1
; GENERAL INFORMATION:
; APPLICANT: Cox III, George N
; APPLICANT: Bolder Biotechnology, Inc.
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4152-1-PUS
; CURRENT APPLICATION NUMBER: US/10/400,708
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US/09/462,941
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-400-708-6

```
Query Match        100.0%;   Score 896;  DB 14;   Length 174;
Best Local Similarity 100.0%;   Pred. No. 2.3e-85;
Matches 174; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120

Qy  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
```

Search completed: October 6, 2004, 16:18:30
Job time : 128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 15:59:42 ; Search time 39 Seconds
(without alignments)
429.162 Million cell updates/sec

Title: US-10-009-792C-19
Perfect score: 896
Sequence: 1 TPLGPASLPQSPLLKCLEQ........SHLQSFLEVSYRVLRHLAQP 174

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	896	100.0	204	1	FQHUGL	granulocyte colony
2	884.5	98.7	207	2	A24573	granulocyte colony
3	744	83.0	194	2	T09255	granulocyte colony
4	738	82.4	174	2	T10268	granulocyte colony
5	641	71.5	208	2	A26496	granulocyte colony
6	634	70.8	214	2	JC5043	granulocyte colony
7	304	33.9	201	2	A42247	myelomonocytic gro
8	106.5	11.9	212	2	I46621	prointerleukin 6 -
9	106.5	11.9	212	2	I46590	interleukin 6 - pi
10	100	11.2	208	2	T09216	interleukin-6 prec
11	94	10.5	345	2	C82270	hypothetical prote
12	89.5	10.0	2175	1	S03170	homeotic protein c
13	88.5	9.9	666	2	A87577	oligopeptide trans
14	88.5	9.9	786	2	F83292	probable sensor/re
15	86	9.6	208	1	A56610	interleukin-6 prec
16	86	9.6	502	2	S46506	H+-transporting tw
17	85	9.5	207	2	I46084	interleukin 6 - ca
18	84	9.4	274	1	C69362	conserved hypothet
19	83.5	9.3	846	2	JC7721	aryl hydrocarbon r
20	80.5	9.0	406	2	B72766	probable threonyl-
21	80	8.9	474	2	D75550	probable D-alanyl-
22	80	8.9	477	2	T46304	hypothetical prote
23	79.5	8.9	351	2	T19623	hypothetical prote
24	79	8.8	423	2	AC3553	4-aminobutyrate tr
25	78	8.7	208	1	S29549	interleukin-6 - sh
26	78	8.7	316	2	H82958	homoserine kinase
27	77.5	8.6	653	2	C82580	oligopeptide trans
28	77	8.6	974	2	AC2076	two-component hybr
29	76.5	8.5	195	2	JH0680	ciliary neurotroph
30	76.5	8.5	506	2	T07942	probable squalene
31	76	8.5	469	2	AD1926	hypothetical prote
32	76	8.5	1704	2	S71363	probable ATP-bindi
33	76	8.5	1704	2	A59188	ATP-binding casset
34	75.5	8.4	156	2	JQ1138	rRNA methyltransfe
35	75.5	8.4	3027	2	JQ1917	polyprotein-pars
36	75	8.4	199	2	B38285	interleukin-11 pre
37	75	8.4	219	2	C75637	hypothetical prote
38	75	8.4	245	2	E70583	hypothetical prote
39	75	8.4	472	2	A70951	hypothetical prote
40	74.5	8.3	212	1	IVHUB2	interleukin-6 prec
41	74.5	8.3	254	2	S34724	probable oxidoredu
42	74.5	8.3	920	2	JC7313	aryl hydrocarbon r
43	74	8.3	367	2	I39172	cyclin-dependent k
44	74	8.3	453	2	H87373	hypothetical prote
45	74	8.3	638	2	A36929	virulence regulato

ALIGNMENTS

RESULT 1
FQHUGL
granulocyte colony-stimulating factor precursor - human
N;Alternate names: colony-stimulating factor 3; G-CSF
C;Species: Homo sapien (man)
C;Date: 30-Jun-1988 #sequence_revision 18-Aug-1995 #text_change 22-Jun-1999
C;Accession: A25093; A49796; A47587; S68331
R;Nagata, S.; Tsuchiya, M.; Asano, S.; Yamamoto, O.; Hirata, Y.; Kubota, N.; Oheda, M.;
EMBO J. 5, 575-581, 1986
A;Title: The chromosomal gene structure and two mRNAs for human granulocyte colony-stim
A;Reference number: A25093; MUID:86220137; PMID:2423327
A;Accession: A25093
A;Molecule type: DNA; mRNA
A;Residues: 1-204 <NAG>
A;Cross-references: EMBL:X03656; EMBL:X03655; NID:g31693; PIDN:CAA27290.1; PID:g732764
R;Devlin, J.J.; Devlin, P.E.; Myambo, K.; Lilly, M.B.; Rado, T.A.; Warren, M.K.
J. Leukoc. Biol. 41, 302-306, 1987
A;Title: Expression of granulocyte colony-stimulating factor by human cell lines.
A;Reference number: A49796; MUID:87196936; PMID:3494801
A;Accession: A49796
A;Molecule type: mRNA
A;Residues: 1-204 <DEV>
A;Cross-references: GB:M17706; NID:g183040; PIDN:AAA35882.1; PID:g183041
R;Souza, L.M.; Boone, T.C.; Gabrilove, J.; Lai, P.H.; Zsebo, K.M.; Murdock, D.C.; Chazi
Science 232, 61-65, 1986
A;Title: Recombinant human granulocyte colony-stimulating factor: effects on normal and
A;Reference number: A47587; MUID:86151684; PMID:2420009
A;Accession: A47587
A;Molecule type: mRNA
A;Residues: 19-204 <SOU>
A;Cross-references: GB:M13008; NID:g183044; PIDN:AAA03056.1; PID:g183045
R;Haniu, M.; Horan, T.; Arakawa, T.; Le, J.; Katta, V.; Rohde, M.F.
Arch. Biochem. Biophys. 324, 344-356, 1995
A;Title: Extracellular domain of granulocyte-colony stimulating factor receptor.
A;Reference number: S68331; MUID:96132662; PMID:8554326
A;Accession: S68331
A;Molecule type: protein
A;Residues: 'M',31-46 <HAN>
C;Genetics:
A;Gene: GDB:CSF3
A;Cross-references: GDB:119083; OMIM:138970
A;Map position: 17q11.2-17q12
A;Introns: 14/1; 65/3; 101/3; 150/3
C;Function:
A;Description: stimulates the differentiation and proliferation of hematopoietic progen
C;Superfamily: interleukin-6
C;Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-204/Product: granulocyte colony-stimulating factor #status predicted <MAT>
F;66-72,94-104/Disulfide bonds: #status predicted

```
Query Match          100.0%;    Score 896; DB 1;   Length 204;
Best Local Similarity  100.0%;   Pred. No. 7.8e-78;
Matches 174;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   31  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP   90

Qy   61  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ  150

Qy  121  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP       174
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  151  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP       204
```

RESULT 2
A24573
granulocyte colony-stimulating factor precursor variant splice form CHU-2 - human
N;Alternate names: colony-stimulating factor 3; G-CSF
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C;Accession: A24573
R;Nagata, S.; Tsuchiya, M.; Asano, S.; Kaziro, Y.; Yamazaki, T.; Yamamoto, O.; Hirata, Y
Nature 319, 415-418, 1986
A;Title: Molecular cloning and expression of cDNA for human granulocyte colony-stimulati
A;Reference number: A24573; MUID:86118679; PMID:3484805
A;Accession: A24573
A;Molecule type: mRNA
A;Residues: 1-207 <NAG>
A;Cross-references: EMBL:X03438; NID:g31689; PIDN:CAA27168.1; PID:g31690
C;Comment: This variant splice form is not expressed in three other cell lines and may
C;Genetics:
A;Gene: GDB:CSF3
A;Cross-references: GDB:119083; OMIM:138970
A;Map position: 17q11.2-17q12
A;Introns: 14/1; 68/3; 104/3; 153/3
C;Superfamily: interleukin-6
C;Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-207/Product: granulocyte colony-stimulating factor variant splice form CHU-2 #statu
F;69-75,97-107/Disulfide bonds: #status predicted

```
Query Match          98.7%;     Score 884.5; DB 2;  Length 207;
Best Local Similarity  98.3%;    Pred. No. 9.9e-77;
Matches 174;  Conservative 0;  Mismatches 0;  Indels 3;  Gaps 1;

Qy    1  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKL---CATYKLCHPEELVLLGHSLGIP   57
         |||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db   31  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLVSECATYKLCHPEELVLLGHSLGIP   90

Qy   58  WAPLSSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTI  117
         |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91  WAPL-SSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTI  150

Qy  118  WQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP     174
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  151  WQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP     207
```

RESULT 3
T09255
granulocyte colony-stimulating factor precursor - cat (fragment)
C;Species: Felis silvestris catus (domestic cat)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 20-Jun-2000
R;Dunham, S.P.; Onions, D.E.
submitted to the EMBL Data Library, September 1996
A;Description: Cloning, sequence and expression of feline granulocyte colony stimulating
A;Reference number: Z16630
A;Accession: T09255
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA
A;Residues: 1-194 <DUN>
A;Cross-references: EMBL:Y08558
C;Function:
A;Description: stimulates the differentiation and proliferation of hematopoietic progeni
C;Superfamily: interleukin-6
C;Keywords: cytokine; growth factor; macrophage
F;1-20/Domain: signal sequence (fragment) #status predicted <SIG>
F;21-194/Product: granulocyte colony-stimulating factor #status predicted <MAT>

```
Query Match          83.0%;     Score 744; DB 2;  Length 194;
Best Local Similarity  81.0%;    Pred. No. 2.1e-63;
Matches 141;  Conservative 11;  Mismatches 22;  Indels 0;  Gaps 0;

Qy    1  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP   60
         ||||| ||||||||||||||||||||||||||||||||||||||| |:|||||||:|||
Db   21  TPLGPTSSLPQSFLLKCLEQVRKVQADGTALQERLCAAHKLCHPEELVLLGHALGIPQAP   80

Qy   61  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ  120
         ||||||||| ||| |:||||| || | |||||:|||| ||| :|||:||  || ||| :
Db   81  LSSCSSQALQLTGCLRQLHSGLFLYQGLLQALAGISPELAPTLDMLQLDITDFAINIWQQ  140

Qy  121  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP       174
         |:::|||| |:|| ||||||::||:||||||||||:::|||||| :|| : ||:|
Db  141  MEDVGMAPAVPPTQGTMPTTSAFQRRAGGTLVASNLQSFLEVAYRALRHFTKP       194
```

RESULT 4
T10268
granulocyte colony-stimulating factor - sheep (fragment)
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T10268
R;O'Brien, P.M.; Seow, H.F.; Rothel, J.S.; Wood, P.R.
DNA Seq. 4, 339-342, 1994
A;Title: Cloning and sequencing of an ovine granulocyte colony-stimulating factor cDNA.
A;Reference number: Z17009; MUID:95102116; PMID:7528579
A;Accession: T10268
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-174 <OBR>
A;Cross-references: EMBL:L07939; NID:g310381; PIDN:AAA68006.1; PID:g310382
C;Genetics:
A;Gene: CSF
C;Function:
A;Description: stimulates the differentiation and proliferation of hematopoietic progeni
C;Superfamily: interleukin-6
C;Keywords: cytokine; growth factor; macrophage

```
Query Match          82.4%;     Score 738; DB 2;  Length 174;
Best Local Similarity  82.2%;    Pred. No. 6.9e-63;
Matches 143;  Conservative 10;  Mismatches 21;  Indels 0;  Gaps 0;

Qy    1  TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP   60
         |||||| ||||||||||||||||| |:||| |:|||||::|||||||||:||||||| |||
Db    1  TPLGPARSLPQSFLLKCLEQVRKIQADGAELQERLCATHKLCHPEELVLLGHSLGIPQAP   60

Qy   61  LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ  120
         ||||  |||||| |||||| |||||||||||:|||| ||| :|||:|| || :: |||
Db   61  LSSCSSQSLQLTSCLDQLHGGLFLYQGLLQALAGISPELAPTLDVTDFATNIWLQ       120

Qy  121  MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP       174
         |:::|||| |:||:| ||||::|:||||||:|||||:::|||||| :|| : ||:|
Db  121  MEDLGVAPAVQPTQGTMPTTSAFQRRAGGVLVASQLQRFLGLAYRGRYLAEP        174
```

RESULT 5
A26496
granulocyte colony-stimulating factor precursor - mouse
N;Alternate names: G-CSF
C;Species: Mus musculus (house mouse)
C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 16-Jul-1999
C;Accession: A26496; A26496; S02493

R:Tsuchiya, M.; Kaziro, Y.; Nagata, S.
Eur. J. Biochem. 165, 7-12, 1987
A;Title: The chromosomal gene structure for murine granulocyte colony-stimulating factor
A;Reference number: A29536
A;Accession: A29536
A;Molecule type: DNA
A;Residues: 1-208 <TSU>
A;Cross-references: GB:X05402; NID:g51059; PIDN:CAA28996.1; PID:g51060
R:Tsuchiya, M.; Asano, S.; Kaziro, Y.; Nagata, S.
Proc. Natl. Acad. Sci. U.S.A. 83, 7633-7637, 1986
A;Title: Isolation and characterization of the cDNA for murine granulocyte colony-stimul
A;Reference number: A26496; MUID:87017003; PMID:3489940
A;Accession: A26496
A;Molecule type: mRNA
A;Residues: 1-208 <TS2>
A;Cross-references: GB:M13926; NID:g193451; PIDN:AAA37672.1; PID:g309248
R:Simpson, R.J.; Nice, E.C.; Nicola, N.A.
Biol. Chem. Hoppe-Seyler 368, 1327-1331, 1987
A;Title: Structural studies on the murine granulocyte colony-stimulating factor.
A;Reference number: S02493; MUID:88106998; PMID:3501294
A;Accession: S02493
A;Status: preliminary
A;Molecule type: protein
A;Residues: 31-34,'X',36-43;48-51,'X',53-57;'X',60-71;159-164,'X',166-176;183-198,'X',20
C;Genetics:
A;Introns: 14/1;71/3;107/3;156/3
C;Superfamily: interleukin-6
C;Keywords: cytokine; growth factor; macrophage; monomer

Query Match   71.5%; Score 641; DB 2; Length 208;
Best Local Similarity 75.9%; Pred. No. 1.4e-53;
Matches 129; Conservative 9; Mismatches 32; Indels 0; Gaps 0;

Qy   3 LGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLS 62
          |:||||||||||:||||||||| :||||||||||||||||||||||||:||||||
Db   39 LPPSLPLPRSFLLKSLEQVRKIQASGSVLLEQLCATYKLCHPEELVLLGHSLGIPKASLS 98

Qy   63 SCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQME 122
          ||||||| ||||||||||| |||||||||| ||:|| :||||:|||||||:||||||||||
Db   99 GCSSQALQQTQCLSQLHSGLCLYQGLLQALSGISPALAPTLDLLQLDVANFATTIWQQME 158

Qy 123 ELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLA 172
          :||:||:|||||:|||||:|||||||||||| ||:|||||| ||||| ||
Db 159 NLGVAPTVQPTQSAMPAFTSAFQRRAGGVLAISYLQGFLETARLALHHLA 208

RESULT 6
JC5043
granulocyte colony-stimulating factor precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 16-Jul-1999
C;Accession: JC5043
R;Han, S.W.; Ramesh, N.; Osborne, W.R.A.
Gene 175, 101-104, 1996
A;Title: Cloning and expression of the cDNA encoding rat granulocyte colony-stimulating
A;Reference number: JC5043; MUID:97074656; PMID:8917083
A;Accession: JC5043
A;Molecule type: mRNA
A;Residues: 1-214 <HAN>
A;Cross-references: GB:U37101; NID:g1680658; PIDN:AAC52915.1; PID:g1680659
A;Experimental source: skin fibroblasts
C;Comment: This receptor acts on precursor hemapoietic cells to control the production a
C;Superfamily: interleukin-6
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-214/Product: granulocyte colony-stimulating factor #status predicted <MAT>

Query Match   70.8%; Score 634; DB 2; Length 214;
Best Local Similarity 73.8%; Pred. No. 6.9e-53;
Matches 127; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

Qy   3 LGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLS 62
          |:|||||||||||:||||||||| :||||||||||||||||||||:||||||:||||||
Db   30 LPPSLPLPRSFLLKSLEQVRKIQARNTELLEQLCATYKLCHPEELVLFGHSLGIPKASLS 89

Qy   63 SCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQME 122
          ||:|||| | |||||||||||||||||||| || :|||| | ||||:||||||||||
Db   90 SCSSQALQQTKCLSQLHSGLFLYQGLLQALALAGISSLAPTLMLHLDVDNFATTIWQQME 149

Qy 123 ELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
          :||:||:|||||: ||:|:|||||||||||||:| |||||| : ||||| |
Db 150 SLGVAPTVQPTQSTMPIFTSAFQRRAGGVLVTSYLQSFLETAHHALHHLPRP 201

RESULT 7
A42247
myelomonocytic growth factor precursor - chicken
N;Alternate names: colony-stimulating factor cMGF
C;Species: Gallus gallus (chicken)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C;Accession: A42247; S03633
R;Sterneck, E.; Blattner, C.; Graf, T.; Leutz, A.
Mol. Cell. Biol. 12, 1728-1735, 1992
A;Title: Structure of the chicken myelomonocytic growth factor gene and specific activat
A;Reference number: A42247; MUID:9219519; PMID:1549124
A;Accession: A42247
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-201 <STE>
A;Note: sequence extracted from NCBI backbone (NCBIN:89832, NCBIP:89836)
R;Leutz, A.; Damm, K.; Sterneck, E.; Kowenz, E.; Ness, S.; Frank, R.; Gausepohl, H.; Pa
EMBO J. 8, 175-181, 1989
A;Title: Molecular cloning of the chicken myelomonocytic growth factor (cMGF) reveals r
A;Reference number: S03633; MUID:89231616; PMID:2785450
A;Accession: S03633
A;Molecule type: mRNA
A;Residues: 1-201 <LEU>
A;Cross-references: EMBL:X14477; NID:g63596; PIDN:CAA32639.1; PID:g63597
C;Superfamily: interleukin-6
C;Keywords: glycoprotein
F;1-23/Domain: signal sequence predicted <SIG>
F;24-201/Product: myelomonocytic growth factor #status predicted <MAT>
F;123,137/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match   33.9%; Score 304; DB 2; Length 201;
Best Local Similarity 40.6%; Pred. No. 1.6e-21;
Matches 67; Conservative 28; Mismatches 68; Indels 2; Gaps 1;

Qy   11 QSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQ 70
          | ||:| || :||| || ||||: | |: ||| || || |:|| || : ||:
Db   36 QLFLHKNLEFTRKIRGDVAALQRAVCDTFQLCTEEELQLVQPDPHLVQAPLDQCHKRGFQ 95

Qy   71 LAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPAL 130
           | | : | || | :||| || | | :||| |:: : :|:: :|| :
Db   96 AEVCFTQIRAGLHAVHDSLGAVLRLLPNHTTIVETLQLDAANLSSNIQQOMEDLGLDTVT 155

Qy 131 QPTQ--GAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQ 173
          :| : | | | | : || |: | |||| : | ||||:
Db 156 LPAEQRSPPPTFSGPFQQQVGGFFILANFQRFLETAYRALRHLAR 200

RESULT 8
I46621
prointerleukin 6 - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C;Accession: I46621
R;Richards, C.; Saklatva, J.
Cytokine 3, 269-276, 1991
A;Title: Molecular Cloning and Sequence of Porcine Interleukin 6 cDNA and Expression of
A;Reference number: I46621; MUID:91338547; PMID:1873476
A;Accession: I46621
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-212 <RIC>
A;Cross-references: GB:M86722; NID:g164624, PIDN:AAC37333.1; PID:g164625
C;Genetics:

A;Gene: IL6
C;Superfamily: interleukin-6

C;Keywords: cytokine; growth factor

```
Query Match        11.2%;    Score 100; DB 2;  Length 208;
Best Local Similarity   21.6%;  Pred. No. 0.041;
Matches   38;  Conservative   35;  Mismatches   95;  Indels   8;  Gaps   3;

Qy      2 PLGPASSLPQSFLL----KCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP   57
                  ::  : :  :   ::::     :|||||:||| :|| |||:||||||||||||||||||
Db     31 PLGDETTSNGPLLTADKTKQHIKYILGKISALKNEMCNNFSKCENSKEVLAENNLNLP     90

Qy     58 -WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATT  116
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     91 KMAEKDGCFQSGFNQETCLMKITTGLSEFQIYLEVLQNEFKGEKENIKTMQISTKVLVQI  150

Qy    117 IWQQMEELGMA---PALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR     169
              : ::  :      ::      :     :        ::: :| |: ||||||||
Db    151 LMQKMKNPEVTTPDTAKSSLLAKLHSQNEWLKNTTHLILRSLEDFIQFSLRAVR      206
```

RESULT 11
C82270
hypothetical protein VC0886 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: C82270
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
l., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: C82270
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-345 <HEI>
A;Cross-references: GB:AE003852; NID:g9655323; PIDN:AAF94048.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0886
A;Map position: 1

```
Query Match        10.5%;    Score 94; DB 2;  Length 345;
Best Local Similarity  -32.2%;  Pred. No. 0.28;
Matches   38;  Conservative   15;  Mismatches   43;  Indels   22;  Gaps   6;

Qy      7 SSLPQSFLKCLEQVRKIQGDGAALQEKLC---ATYKLCHPEELVLLGHSLGIPWAPLS    62
                 :::    :  : :: :     :    :||||||||||||||||||||
Db    228 ASDPDLFLIAAL--VRALAGAPANILDATCDAILASPAICHPEVLIALA--GRSWMSLE   282

Qy     63 S---CPSQALQAGCLSQ-LHSGLFLYQGLLQALEGI--------SPELGPTLDTLQ     107
                :|           ::: :|| :|||:| :|
Db    283 NSGRAQRFLIRLAQTQNQPLFNQLFADVVMLPALRGVMLPLLHASPSPELAQALEQLQ   340
```

RESULT 12
S03170
homeotic protein cut - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S03170
R;Blochlinger, K.; Bodmer, R.; Jack, J.; Jan, L.Y.; Jan, Y.N.
Nature 333, 629-635, 1988
A;Title: Primary structure and expression of a product from cut, a locus involved in spe
A;Reference number: S03170; MUID:88232956; PMID:2897632
A;Accession: S03170
A;Molecule type: mRNA
A;Residues: 1-2175 <BLO>
A;Cross-references: EMBL:X07985; NID:g7767; PIDN:CAA30794.1; PID:g7768
C;Genetics:
A;Gene: cut
A;Cross-references: FlyBase:FBgn0004198
C;Superfamily: homeotic protein cut; cut repeat homology; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

---

A;Gene: IL6
C;Superfamily: interleukin-6

```
Query Match        11.9%;    Score 106.5; DB 2;  Length 212;
Best Local Similarity   22.6%;  Pred. No. 0.01;
Matches   36;  Conservative   33;  Mismatches   85;  Indels   5;  Gaps   3;

Qy     16 KCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP-WAPLSSCPSQALQLAGC   74
              :||||:||| :|| |||:||||||||||||||||||||||| ||||||||||||||||
Db     52 KTEELIKYILGKISAMRKEMCEKYEKCENSKEVLAENNLNLPKMAEKDGCFQSGFNQETC  111

Qy     75 LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPT-  133
           |||||||||||||||||||||||||||||||||||| |||:|:||:||| |||  :  |
Db    112 LMRITTGLVEFQIYLDYLQKEYESNKGNVEAVQISTKALIQTLRQKGKNPDKATTPNPTT  171

Qy    134 -QGAMPAFAS--AFQRRAGGVLVASHLQSFLEVSYRVLR                       169
               : :|:     :|  ::|  : |: ||  ::: |:
Db    172 NAGLLDKLQSQNEWMKNTKILILRSLEDFLQFSLRAIR                        210
```

RESULT 9
I46590
interleukin 6 - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C;Accession: I46590
R;Mathialagan, N.; Bixby, J.; Roberts, M.R.
Mol. Reprod. Dev. 32, 324-330, 1992
A;Title: Expression of interleukin-6 in porcine, ovine, and bovine preimplantation conce
A;Reference number: I46590; MUID:92360284; PMID:1497880
A;Accession: I46590
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-212 <MAT>
A;Cross-references: GB:M80258; NID:g164514; PIDN:AAC27127.1; PID:g164515
C;Genetics:
A;Gene: IL-6
C;Superfamily: interleukin-6

```
Query Match        11.9%;    Score 106.5; DB 2;  Length 212;
Best Local Similarity   22.6%;  Pred. No. 0.01;
Matches   36;  Conservative   33;  Mismatches   85;  Indels   5;  Gaps   3;

Qy     16 KCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP-WAPLSSCPSQALQLAGC   74
              :||||:||| :|| |||:||||||||||||||||||||||| ||||||||||||||||
Db     52 KTEELIKYILGKISAMRKEMCEKYEKCENSKEVLAENNLNLPKMAEKDGCFQSGFNQETC  111

Qy     75 LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPT-  133
           |||||||||||||||||||||||||||||||||||| |||:|:||:||| |||  :  |
Db    112 LMRITTGLVEFQIYLDYLQKEYESNKGNVEAVQISTKALIQTLRQKGKNPDKATTPNPTT  171

Qy    134 -QGAMPAFAS--AFQRRAGGVLVASHLQSFLEVSYRVLR                       169
               : :|:     :|  ::|  : |: ||  ::: |:
Db    172 NAGLLDKLQSQNEWMKNTKILILRSLEDFLQFSLRAIR                        210
```

RESULT 10
T09216
interleukin-6 precursor - horse
C;Species: Equus caballus (domestic horse)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999
C;Accession: T09216
R;Swiderski, C.B.; Horohov, D.W.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z16613
A;Accession: T09216
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-208 <SWI>
A;Cross-references: EMBL:U64794; NID:g2654387; PIDN:U64794; NID:g2654388
C;Genetics:
A;Gene: IL-6
C;Superfamily: interleukin-6

F;886-958/Domain: cut repeat homology <CU1>
F;1339-1411/Domain: cut repeat homology <CU2>
F;1617-1689/Domain: cut repeat homology <CU3>
F;1746-1802/Domain: homeobox homology <HOX>

Query Match   10.0%; Score 89.5; DB 1; Length 2175;
Best Local Similarity   23.8%; Pred. No. 6.9;
Matches   35; Conservative 20; Mismatches 61; Indels 31; Gaps 3;

Qy   27 DGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAGCLSQLHSGLFLYQ 86
             :||:::| ||||:: || ||:
Db   1404 DENAVHKLVASQYKIA-PEKLMRTGSYSGSPQMP------------------------Q 1437

Qy   87 GLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPTQGAMPAFASAFQR 146
             |:| : |:| ||:
Db   1438 GLASKMQAASLPMQKMMSELKLQEPAQAQHLMQQMQAAAMSAAMQQQQ-----VAQAQQQ 1492

Qy   147 RAGGVLVASHLQSFLEVSYRVLRHLAQ 173
             ||||
Db   1493 AQQAQQAQQHLQQQAQQHLQQQQQHLAQ 1519

RESULT 13
A87577
oligopeptide transporter, OPT family CC2646 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: A87577
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
  B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor
  n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
  Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87577
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-666 <STO>
A;Cross-references: GB;AE005673; NID:g13424225; PIDN:AAK24613.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC2646
C;Superfamily: Pyrococcus horikoshii hypothetical protein PH0361

Query Match   9.9%; Score 88.5; DB 2; Length 666;
Best Local Similarity   32.5%; Pred. No. 2.1;
Matches   37; Conservative 14; Mismatches 48; Indels 15; Gaps 4;

Qy   50 LGHSLGIPWAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLD 109
             | | ||| :::|| | |: :|||||
Db   196 LGAASGATGIGASS--SLALMGAGHLMGITVGVAMFTGLFIAWAILVPILTLVTPMEAD 253

Qy   110 VADFATTIWQ-QMEELG--------MAPALQ-TCGAMPAFASAFQRRAGG 150
             ||| ||| : :| :| |
Db   254 AATHALTVWKSQVRFLGAGVIGAAAIWTLAKLVGPITSGLKSAFAAAQARKAGG 307

RESULT 14
F83292
probable sensor/response regulator hybrid PA2824 [imported] - Pseudomonas aeruginosa (st
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: F83292
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
  adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
  .; Lory, S.; Olson, M.V.
  Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83292
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-786 <STO>
A;Cross-references: GB;AE004709; GB;AE004091; NID:g9948904; PIDN:AAG06212.1; GSPDB:GN001

A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2824

Query Match   9.9%; Score 88.5; DB 2; Length 786;
Best Local Similarity   29.2%; Pred. No. 2.5;
Matches   50; Conservative 19; Mismatches 49; Indels 53; Gaps 10;

Qy   26 GDGAALQEKLC-ATYKLCHPEELVLLG--HSLGIPWAPLSSCPSQALQLAG-CLSQLHSG 81
             | |:|::| ::| || | :| |::| | || | :
Db   490 GLGLALTRKLCEAMQGELTVESTVGLGLFSVGLLAPVSP-PLQALPLRGRVIAQCSAN 548

Qy   82 LFLYQGLLQALEGISPE-----------LGPTLDTLQLDVADFATTIWQQMEELGMA 127
             |:|| |:| ::||
Db   549 ----SGLAQLLQTWLPRWGLEYKRLETDDSLLGHSLDVLISDCPDCL------MGLR 595

Qy   128 PALQPTQGAMPAFASAFQRRAGGVLVASHLQSFL-EVSYRV--LRHLAQP 174
             |||| | | || :: :| :|||| :   ::|:
Db   596 PSI----------------------GTPILLVTAYGSFLEPELARRLSPLRQLARP 629

RESULT 15
A56610
interleukin-6 precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A56610; S22162
R;Droogmans, L.; Cludts, I.; Cleuter, Y.; Kettmann, R.; Burny, A.
  DNA Seq. 2, 411-413, 1992
A;Title: Nucleotide sequence of bovine interleukin-6 cDNA.
A;Reference number: A56610; MUID:93076003; PMID:1446077
A;Accession: A56610
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-208 <DRO>
A;Cross-references: EMBL;X57317; NID:g2193; PIDN:CAA40572.1; PID:g2194
A;Experimental source: BLV induced B cell-lymphosarcoma
A;Note: sequence extracted from NCBI backbone (NCBIP:118917)
C;Superfamily: interleukin-6
C;Keywords: cytokine

Query Match   9.6%; Score 86; DB 1; Length 208;
Best Local Similarity   20.5%; Pred. No. 0.89;
Matches   36; Conservative 37; Mismatches 95; Indels 8; Gaps 3;

Qy   2 PLG---PASSLPQSFLLKCLEQ---VRKIQGDGAALQEKLCATYKLCHPEELVLLGHSL 54
             |:   :| :| || | :|| | |: :
Db   31 PLGEDFKNDTTPGRLLITPEKTEALIKRMVDKISAMRKEICEKNDECESSKETLAENKL 90

Qy   55 GIP-WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADF 113
             || ||| | :
Db   91 NLPKMEEKDGCFQSGFNQAICLIRITAGLLEYQIYLDYLQNEYEGNQENVRDLRKNIRTL 150

Qy   114 ATTIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR 169
             : ::: :   : :|
Db   151 IQILKQKIADLITTPATNTDLLEKMQSSNEWVKNAKILILIRNLENFLQFSLRAIR 206

Search completed: October 6, 2004, 16:07:18
Job time : 43 secs

This Page Blank (uspto)

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:   October 6, 2004, 15:51:18 ; Search time 23 Seconds
                                      (without alignments)
                        393.922 Million cell updates/sec

Title:          US-10-009-792C-19
Perfect score:  896
Sequence:       1 TPLGPASSLPQSFLLKCLEQ.........SHLQSFLEVSYRVLRHLAQP 174

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:   141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      SwissProt_42:*

   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	884.5	98.7	207	1	CSF3_HUMAN	P09919 homo sapien
2	744	83.0	194	1	CSF3_FELCA	O02708 felis silve
3	738	82.4	174	1	CSF3_SHEEP	Q28746 ovis aries
4	734	81.9	175	1	CSF3_CANFA	P35834 canis famil
5	727	81.1	195	1	CSF3_BOVIN	P35833 bos taurus
6	702	78.3	195	1	CSF3_PIG	O02837 sus scrofa
7	641	71.5	208	1	CSF3_MOUSE	P09920 mus musculu
8	304	33.9	201	1	MGF_CHICK	P13854 gallus gall
9	106.5	11.9	212	1	IL6_PIG	P26893 sus scrofa
10	101	11.3	205	1	IL6_ORCOR	Q28747 orcinus orc
11	100	11.2	208	1	IL6_HORSE	Q95181 equus cabal
12	95	10.6	208	1	IL6_FELCA	P41683 felis silve
13	89.5	10.0	2175	1	HMCU_DROME	P10180 drosophila
14	86	9.6	208	1	IL6_BOVIN	P26892 bos taurus
15	84.5	9.4	209	1	IL6_PHOVI	Q28819 phoca vitul
16	84	9.4	208	1	IL6_CAPHI	Q28319 capra hircu
17	84	9.4	208	1	IL6_SHEEP	Q29455 ovis aries
18	80.5	9.0	406	1	SYT2_AERPE	Q9yfy3 aeropyrum p
19	79.5	8.9	207	1	IL6_CANFA	P41323 canis famil
20	78.5	8.7	300	1	GGPP_MOUSE	Q9wtn0 m geranylge
21	78	8.7	316	1	KHSE_PSEAE	Q29364 pseudomonas
22	77	8.6	247	1	RECO_NEIGO	Q9zhy2 neisseria g
23	76.5	8.5	195	1	CNTF_CHICK	Q02011 gallus gall
24	76.5	8.5	294	1	GGPP_BOVIN	P56966 b geranylge
25	76.5	8.5	383	1	ARGE_SHIFL	P59600 shigella fl
26	76.5	8.5	506	1	ER11_BRANA	O65727 brassica na
27	76	8.5	1704	1	ABC3_HUMAN	Q99758 homo sapien
28	75.5	8.4	156	1	KAMC_SACHI	P25919 saccharopol
29	75.5	8.4	3027	1	POLG_PYFV1	Q05057 parsnip yel
30	75	8.4	199	1	IL11_HUMAN	P20809 homo sapien
31	75	8.4	330	1	OXF2_HUMAN	Q8nh61 homo sapien
32	74.5	8.3	212	1	IL6_HUMAN	P05231 homo sapien
33	74	8.3	367	1	CD5S_HUMAN	Q13319 homo sapien
34	74	8.3	2504	1	FAS_HUMAN	P49327 homo sapien
35	73.5	8.2	383	1	ARGE_ECOLI	P23908 escherichia
36	73.5	8.2	737	1	RECG_MYCTU	P95122 mycobacteri
37	73	8.1	204	1	CTF2_MOUSE	P83714 mus musculu
38	73	8.1	836	1	VG26_BPML5	Q05233 mycobacteri
39	72.5	8.1	203	1	Y626_VIBVU	P59257 vibrio vuln
40	72.5	8.1	383	1	ARGE_ECO57	Q8x742 escherichia
41	72.5	8.1	383	1	ARGE_SCOL6	Q8fb97 escherichia
42	72	8.0	1399	1	Z291_HUMAN	Q9by12 homo sapien
43	71.5	8.0	300	1	GGPP_HUMAN	O95749 h geranylge
44	71.5	8.0	518	1	ER12_BRANA	O65726 brassica na
45	71.5	8.0	836	1	GCSR_HUMAN	Q99062 homo sapien

ALIGNMENTS

```
RESULT 1
CSF3_HUMAN
ID   CSF3_HUMAN     STANDARD;      PRT;   207 AA.
AC   P09919;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Granulocyte colony-stimulating factor precursor (G-CSF) (Pluripoietin)
DE   (Filgrastim) (Lenograstim).
GN   CSF3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86118679; PubMed=3484805;
RA   Nagata S., Tsuchiya M., Asano S., Kaziro Y., Yamazaki T., Yamamoto O.,
RA   Hirata Y., Kubota N., Oheda M., Nomura H., Ono M.;
RT   "Molecular cloning and expression of cDNA for human granulocyte
RT   colony-stimulating factor.";
RL   Nature 319:415-418(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86220137; PubMed=2423327;
RA   Nagata S., Tsuchiya M., Asano S., Yamamoto O., Hirata Y., Kubota N.,
RA   Oheda M., Nomura H., Yamazaki T.;
RT   "The chromosomal gene structure and two mRNAs for human granulocyte
RT   colony-stimulating factor.";
RL   EMBO J. 5:575-581(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87196936; PubMed=3494801;
RA   Devlin J.J., Devlin P.E., Myambo K., Lilly M.B., Rado T.A.,
RA   Warren M.K.;
RT   "Expression of granulocyte colony-stimulating factor by human cell
RT   lines.";
RL   J. Leukoc. Biol. 41:302-306(1987).
RN   [4]
RP   SEQUENCE FROM N.A., AND VARIANTS MET-157 AND THR-174.
RA   Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 19-207 FROM N.A.
RX   MEDLINE=86151684; PubMed=2420009;
RA   Souza L.M., Boone T.C., Gabrilove J., Lai P.H., Zsebo K.M.,
RA   Murdock D.C., Chazin V.R., Bruszewski J., Lu H., Chen K.K.,
RA   Barendt J., Platzer E., Moore M.A.S., Mertelsmann R., Welte K.;
RT   "Recombinant human granulocyte colony-stimulating factor: effects on
RT   normal and leukemic myeloid cells.";
RL   Science 232:61-66(1986).
RN   [6]
RP   CARBOHYDRATE-LINKAGE SITE.
RX   MEDLINE=93293942; PubMed=7685769;
RA   Clogston C.L., Hu S., Boone T.C., Lu H.S.;
```

```
RT   "Glycosidase digestion, electrophoresis and chromatographic analysis
RT   of recombinant human granulocyte colony-stimulating factor glycoforms
RT   produced in Chinese hamster ovary cells.";
RL   J. Chromatogr. A 637:55-62 (1993).
RN   [7]
RP   STRUCTURE BY NMR.
RX   MEDLINE=93106200; PubMed=1281794;
RA   Zink T., Ross A., Ambrosius D., Rudolph R., Holak T.A.;
RT   "Secondary structure of human granulocyte colony-stimulating factor
RT   derived from NMR spectroscopy.";
RL   FEBS Lett. 314:435-439(1992).
RN   [8]
RP   STRUCTURE BY NMR.
RX   MEDLINE=94304859; PubMed=7518249;
RA   Zink T., Ross A., Luers K., Cieslar C., Rudolph R., Holak T.A.;
RT   "Structure and dynamics of the human granulocyte colony-stimulating
RT   factor determined by NMR spectroscopy. Loop mobility in a four-helix-
RT   bundle protein.";
RL   Biochemistry 33:8453-8463(1994).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX   MEDLINE=93281718; PubMed=7685117;
RA   Hill C.P., Osslund T.D., Eisenberg D.;
RT   "The structure of granulocyte-colony-stimulating factor and its
RT   relationship to other growth factors.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:5167-5171(1993).
CC   -!- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
CC       CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
CC       DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS
CC       OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES. THIS
CC       CSF INDUCES GRANULOCYTES.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P09919-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=P09919-2; Sequence=VSP_002673;
CC   -!- PTM: O-GLYCAN CONSISTS OF GAL-GALNAC DISACCHARIDE WHICH CAN BE
CC       MODIFIED WITH UP TO TWO SIALIC ACID RESIDUES (DONE IN
CC       RECOMBINANTLY EXPRESSED G-CSF FROM CHO CELLS).
CC   -!- PHARMACEUTICAL: Available under the names Neupogen or Granulokine
CC       (Amgen/Roche) and Granocyte (Rhone-Poulenc). Used to treat
CC       neutropenia (a disorder characterized by an extremely low number
CC       of neutrophils in blood).
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   -!- CAUTION: REF.4 MISQUOTES THE GENE NAME AS "CSF1".
CC   -!- DATABASE: NAME=Neupogen/Granulokine;
CC       NOTE=Clinical information on Neupogen/Granulokine;
CC       WWW="http://www.neupogen.com/".
```

```
DR   EMBL; X03438; CAA27168.1; -.
DR   EMBL; M13008; AAA03056.1; -.
DR   EMBL; X03656; CAA27291.1; -.
DR   EMBL; X03655; CAA27290.1; -.
DR   EMBL; AF388025; AAK62469.1; -.
DR   EMBL; M17706; AAA35882.1; -.
DR   PIR; A24573; A24573.
DR   PIR; A25093; FQHUGL.
DR   PDB; 1RHG; 31-JAN-94.
DR   PDB; 1GNC; 31-JUL-94.
DR   PDB; 1CD9; 08-MAR-00.
DR   PDB; 1PGR; 08-MAR-00.
DR   Genew; HGNC:2438; CSF3.

DR   MIM; 138970; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005130; F:granulocyte colony-stimulating factor recep..; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Cytokine; Growth factor; Glycoprotein; Signal; Alternative splicing;
KW   Polymorphism; Pharmaceutical; 3D-structure.
FT   SIGNAL        1     30
FT   CHAIN        31    207       GRANULOCYTE COLONY-STIMULATING FACTOR.
FT   DISULFID     69     75
FT   DISULFID     97    107
FT   CARBOHYD    166    166       O-LINKED (GALNAC. . .) (BY SIMILARITY).
FT   VARSPLIC     66     68       Missing (in isoform Short).
FT                                /FTId=VSP_002673.
FT   VARIANT     157    157       L -> M.
FT                                /FTId=VAR_013073.
FT   VARIANT     174    174       A -> T.
FT                                /FTId=VAR_013074.
FT   HELIX        41     65
FT   HELIX        69     71
FT   HELIX        77     86
FT   TURN         87     88
FT   HELIX       105    124
FT   TURN        125    127
FT   TURN        130    132
FT   HELIX       133    156
FT   TURN        157    158
FT   HELIX       176    203
FT   TURN        204    204
SQ   SEQUENCE   207 AA;  22293 MW;  421F635ECC776996 CRC64;

Query Match             98.7%;  Score 884.5;  DB 1;  Length 207;
Best Local Similarity   98.3%;  Pred. No. 3.1e-74;
Matches  174;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps   1;

Qy         1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKL---CATYKLCHPEELVLLGHSLGIP 57
             ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db        31 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLVSECATYKLCHPEELVLLGHSLGIP 90

Qy        58 WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTI 117
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        91 WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTI 150

Qy       118 WQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       151 WQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 207
```

RESULT 2
CSF3_FELCA

```
ID   CSF3_FELCA    STANDARD;    PRT;  194 AA.
AC   O02708;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Granulocyte colony-stimulating factor precursor (G-CSF) (Fragment).
GN   CSF3.
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=European shorthair; TISSUE=Lung;
RX   MEDLINE=21389237; PubMed=11497496;
RA   Dunham S.P., Onions D.E.;
```

RT   "Isolation, nucleotide sequence and expression of a cDNA encoding
RT   feline granulocyte colony-stimulating factor.";
RL   Cytokine 14:347-351(2001).
CC   -!- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
CC   DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS,
CC   CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
CC   OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES. THIS
CC   CSF INDUCES GRANULOCYTES (BY SIMILARITY).
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: O-glycosylated (By similarity).
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; Y08558; CAA69853.1; -.
DR   PIR; T09255; T09255.
DR   HSSP; P35834; 1BGE.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Cytokine; Growth factor; Glycoprotein; Signal.
FT   NON_TER   1   1
FT   SIGNAL   <1   20   POTENTIAL.
FT   CHAIN   21   194   GRANULOCYTE COLONY-STIMULATING FACTOR.
FT   DISULFID   56   62   BY SIMILARITY.
FT   DISULFID   84   94   BY SIMILARITY.
FT   CARBOHYD   153   153   O-LINKED (GALNAC...) (BY SIMILARITY).
SQ   SEQUENCE   194 AA; 21154 MW; F72B7AB3DAE7385E CRC64;

Query Match   83.0%; Score 744; DB 1; Length 194;
Best Local Similarity   81.0%; Pred. No. 2.3e-61;
Matches 141; Conservative 11; Mismatches 22; Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||| ||||||||||||||||||| || ||||| ||||: ||||||||||| |||: ||
Db   21 TPLGPTSSLPQSFLLKCLEQVRKVQADGTALQERLCAAHKLCHPEELVLLGHALGIPQAP 80

Qy   61 LSSCPSOALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          |||| ||||| ||| ||| ||||||||||||| ||||| ||| ||||| :|||:|||||
Db   81 LSSCSSOALQLTGCLRQLHSGLFLYQGLLQALAGISPELAPTLDMLQLDITDFAINIWQQ 140

Qy 121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP   174
          ||::| ||| :|||| || |:::||||| |:|||::|||:|| || : |||: :|
Db 141 MEDVGNMAPAVPPTQGTMPFTFTSAFQRRAAGTLVASNLQSFLEVAYRALRHFTKP   194

RESULT 3
CSF3_SHEEP
ID   CSF3_SHEEP   STANDARD;   PRT;   174 AA.
AC   Q28746;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Granulocyte colony-stimulating factor (G-CSF).
GN   CSF3.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.

RX   MEDLINE=9510216; PubMed=7528579;
RA   O'Brien P.M., Seow H.F., Rothel J.S., Wood P.R.;
RT   "Cloning and sequencing of an ovine granulocyte colony-stimulating
RT   factor cDNA.";
RL   DNA Seq. 4:339-342(1994).
CC   -!- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
CC   CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
CC   DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS,
CC   OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES. THIS
CC   CSF INDUCES GRANULOCYTES (BY SIMILARITY).
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: O-glycosylated (By similarity).
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; L07939; AAA68006.1; -.
DR   PIR; T10268; T10268.
DR   HSSP; P35833; 1BGC.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Cytokine; Growth factor; Glycoprotein.
FT   DISULFID   36   42   BY SIMILARITY.
FT   DISULFID   64   74   BY SIMILARITY.
FT   CARBOHYD   133   133   O-LINKED (GALNAC...) (BY SIMILARITY).
SQ   SEQUENCE   174 AA; 18806 MW; BA5AA8F8D23ACD1E CRC64;

Query Match   82.4%; Score 738; DB 1; Length 174;
Best Local Similarity   82.2%; Pred. No. 7.3e-61;
Matches 143; Conservative 10; Mismatches 21; Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
          ||||| ||||||||||||||||||||||| ||||| ||||: ||||||||||||||||||
Db   1 TPLGPARSLPQSFLLKCLEQVRKIQADGAELQERLCATHKLCHPEELVLLGHSLGIPQAP 60

Qy   61 LSSCPSOALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
          |||| | : ||:|| ||| |||||||||||||| ||||| ||||||||| ||||:||| |
Db   61 LSSCSSOSLQLTSCLDQLHGGLFLYQGLLQALAGISPELAPTLDTLQLDVTDFATNIWLQ 120

Qy 121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP   174
          ||:||:||| :|||| || |:::||||||||||||:||:|| ||: :||: :| :|
Db 121 MEDLGVAPAVQPTQGTMPTFTSAFQRRAGGVLVASQLQRFLGLAYRGLRYLAEP   174

RESULT 4
CSF3_CANFA
ID   CSF3_CANFA   STANDARD;   PRT;   175 AA.
AC   P35834;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Granulocyte colony-stimulating factor (G-CSF).
GN   CSF3.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX   MEDLINE=94076341; PubMed=7504736;
RA   LoveJoy B., Cascio D., Eisenberg D.;

RT   "Crystal structure of canine and bovine granulocyte-colony
RT   stimulating factor (G-CSF).";
RL   J. Mol. Biol. 234:640-653(1993).
CC   -!- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
CC   CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
CC   DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS,
CC   OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES. THIS
CC   CSF INDUCES GRANULOCYTES.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: O-glycosylated.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
DR   PDB; 1BGD; 31-OCT-93.
DR   PDB; 1BGE; 31-OCT-93.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Cytokine; Growth factor; Glycoprotein; 3D-structure.
FT   DISULFID   37   43
FT   DISULFID   65   75
FT   CARBOHYD   134   O-LINKED (GALNAC...) (BY SIMILARITY).
FT   STRAND   10   10
FT   HELIX   12   39
FT   HELIX   45   55
FT   TURN   56   56
FT   HELIX   63   65
FT   TURN   67   69
FT   TURN   72   92
FT   TURN   93   95
FT   TURN   98   100
FT   HELIX   101   125
FT   TURN   126   126
FT   HELIX   144   171
FT   HELIX   172   172
FT   STRAND   173   173
SQ   SEQUENCE   175 AA; 18858 MW; 28C26B24990C6DB3 CRC64;

     Query Match   81.9%; Score 734; DB 1; Length 175;
     Best Local Similarity   80.9%; Pred. No. 1.7e-60;
     Matches 140; Conservative 11; Mismatches 22; Indels   0; Gaps   0;

Qy   2 PLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPL 61
              :||| ||:||||||||||||| ||||||||| ||||:|||:||||||||||||||||| |
Db   3 PLGPTGPLPQSFLLKCLEQMRKVQADGTALQETLCATHQLCHPEELVLLGHALGIPQPPL 62

Qy   62 SSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQM 121
              ||| ||||| ||| |||||||||||| |||||||||| || ||||||| || | ||||||
Db   63 SSCSSQALQLAGCLRQLHSGLFLYQGLLQALAGISPELAPTLDTLQLDTTDFAINIWQQM 122

Qy   122 EELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
              ||:|||||| ||:|| ||::|||||||||||||:|||||:|| :| |||| |
Db   123 EDLGMAPAVPPTQGTMPAFTSAFQRRAGGVLVASNLQSFLELAYRALRHFAKP 175

RESULT 5
CSF3_BOVIN
ID   CSF3_BOVIN   STANDARD;   PRT;   195 AA.
AC   P35833; Q9TV89;
DT   01-JUN-1994 (Rel. 29, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Granulocyte colony-stimulating factor precursor (G-CSF).
GN   CSF3 OR GCSF.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]

RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein;
RA   Heidari M., Kehrli M.E. Jr.;
RT   "Cloning, sequencing, and analysis of cDNA encoding bovine granulocyte
RT   colony stimulating factor.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX   MEDLINE=94076341; PubMed=7504736;
RA   Lovejoy B., Cascio D., Eisenberg D.;
RT   "Crystal structure of canine and bovine granulocyte-colony
RT   stimulating factor (G-CSF).";
RL   J. Mol. Biol. 234:640-653(1993).
CC   -!- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
CC   CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
CC   DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS
CC   OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES. THIS
CC   CSF INDUCES GRANULOCYTES.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: O-glycosylated.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF092533; AAD16102.1; -.
DR   PDB; 1BGC; 31-OCT-93.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Cytokine; Growth factor; Glycoprotein; Signal; 3D-structure.
FT   SIGNAL   1   21
FT   CHAIN   22   195   GRANULOCYTE COLONY-STIMULATING FACTOR.
FT   DISULFID   57   63
FT   DISULFID   85   95
FT   CARBOHYD   154   O-LINKED (GALNAC...) (BY SIMILARITY).
FT   CONFLICT   93   94   TS -> RG (IN REF. 2).
FT   HELIX   32   60
FT   HELIX   65   69
FT   TURN   70   71
FT   HELIX   72   75
FT   TURN   76   76
FT   HELIX   83   85
FT   HELIX   87   89
FT   TURN   92   112
FT   TURN   113   115
FT   TURN   118   120
FT   TURN   121   145
FT   HELIX   164   191
FT   TURN   192   192
SQ   SEQUENCE   195 AA; 21431 MW; 8C06119E4ADFBA73 CRC64;

     Query Match   81.1%; Score 727; DB 1; Length 195;
     Best Local Similarity   81.0%; Pred. No. 8.5e-60;
     Matches 141; Conservative 11; Mismatches 22; Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
              :||||| ||||||||||||||||||||||||||||||:|||:||||||||||||:||| |
Db   22 TPLGPARSLPQSFLLKCLEQVRKIQADGAELQERLCAAHKLCHPEELMLLRHSLGIPQAP 81

Qy   61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
              |||| ||||| ||| |||||||||||||||| ||||||| || ||||||| || |||||
Db   82 LSSCSSQSLQLTSCNLHGGLFLYQGLLQALAGISPELAPTLDTLQLDVTDFATNIWLQ 141

```
Qy     121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
               |||| ||||| ||| |:|||:::|||||||||||||||:|| :::||:|
Db     142 MEDLGAAPAVQPTQGAMPTFTSAFQRRAGGVLVASQLHRFLELAYRGLRYLAEP 195

RESULT 6
CSF3_PIG
ID   CSF3_PIG       STANDARD;       PRT;       195 AA.
AC   O02837; O19180;
DT   15-JUL-1998 (Rel. 36, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Granulocyte colony-stimulating factor precursor (G-CSF).
GN   CSF3.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kulmburg P.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Gloster S.E., Sandeman R.M., Strom A.D.G.;
RT   "Cloning of a cDNA and gene encoding porcine granulocyte-colony
RT   stimulating factor.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
    CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
    DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS
    OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES. THIS
    CSF INDUCES GRANULOCYTES (BY SIMILARITY).
-!- SUBUNIT: Monomer.
-!- SUBCELLULAR LOCATION: Secreted.
-!- PTM: O-glycosylated (By similarity).
-!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; Y10494; CAA71518.1; -.
DR   EMBL; U68482; AAB70701.1; -.
DR   EMBL; U68481; AAB70700.1; -.
DR   HSSP; P35833; 1BGC.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Cytokine; Growth factor; Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    195       GRANULOCYTE COLONY-STIMULATING FACTOR.
FT   DISULFID     57     63       BY SIMILARITY.
FT   DISULFID     85     95       BY SIMILARITY.
FT   CARBOHYD    154    154       O-LINKED (GALNAC...) (BY SIMILARITY).
FT   CONFLICT    123    123       A -> R (IN REF. 1).
SQ   SEQUENCE   195 AA;  21214 MW;  84787F20DB0AEA1C CRC64;

Query Match             78.3%;  Score 702; DB 1;  Length 195;
Best Local Similarity   79.2%;  Pred. No. 1.7e-57;
Matches 137;  Conservative 11;  Mismatches 25;  Indels 0;  Gaps 0;

Qy       2 PLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPL 61
```

```
Db      23 PLGPASSLPQSFLLKCLEQVRKIQADGAELQERLCATHKLCHPQELVLLGHSLGLPQASL 82
               |||||||||||||||||||||| :||||:||:|||:|||||:|||||||||||:||:|
Qy      62 SSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQM 121
               |||:||| |||:|||||| ||| :||||||||| ||||| ||||||||||:|||||||
Db      62 SSCSSQALQLTGCLNQLHGGLVLYQGLLQALAGISPELAPALDILQLDVTDLATNIWLQM 142

Qy     122 EELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
               |:|||||||||||||||||||||||||||||:||||||:|:||:||:|| :|
Db     143 EDLRMAPASLPTQGTVPTFTSAFQRRAGGVLVVSQLQGFLELAYRVLRYLAEP 195

RESULT 7
CSF3_MOUSE
ID   CSF3_MOUSE     STANDARD;       PRT;       208 AA.
AC   P09920;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Granulocyte colony-stimulating factor precursor (G-CSF).
GN   CSF3 OR CSFG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87017003; PubMed=3489940;
RA   Tsuchiya M., Asano S., Kaziro Y., Nagata S.;
RT   "Isolation and characterization of the cDNA for murine granulocyte
RT   colony-stimulating factor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:7633-7637(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87190474; PubMed=3494605;
RA   Tsuchiya M., Kaziro Y., Nagata S.;
RT   "The chromosomal gene structure for murine granulocyte colony-
RT   stimulating factor.";
RL   Eur. J. Biochem. 165:7-12(1987).
-!- FUNCTION: GRANULOCYTE/MACROPHAGE COLONY-STIMULATING FACTORS ARE
    CYTOKINES THAT ACT IN HEMATOPOIESIS BY CONTROLLING THE PRODUCTION,
    DIFFERENTIATION, AND FUNCTION OF 2 RELATED WHITE CELL POPULATIONS
    OF THE BLOOD, THE GRANULOCYTES AND THE MONOCYTES-MACROPHAGES. THIS
    CSF INDUCES GRANULOCYTES.
-!- SUBUNIT: Monomer.
-!- SUBCELLULAR LOCATION: Secreted.
-!- PTM: O-glycosylated (By similarity).
-!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; M13926; AAA37672.1; -.
DR   EMBL; X05402; CAA28986.1; -.
DR   PIR; A29536; A26496.
DR   HSSP; P09919; 1RHG.
DR   MGD; MGI:1339751; Csf3.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Cytokine; Growth factor; Glycoprotein; Signal.
FT   SIGNAL        1     30
FT   CHAIN        31    208       GRANULOCYTE COLONY-STIMULATING FACTOR.
FT   DISULFID     72     78       BY SIMILARITY.
```

```
FT   DISULFID     100    110       BY SIMILARITY.
FT   CARBOHYD     169    169       O-LINKED (GALNAC...) (BY SIMILARITY).
SQ   SEQUENCE     208 AA;  22421 MW;  0BF3622135C906DB CRC64;

Query Match             71.5%;  Score 641;  DB 1;  Length 208;
Best Local Similarity   75.9%;  Pred. No. 7.3e-52;
Matches 129;  Conservative 9;  Mismatches 32;  Indels 0;  Gaps 0;

Qy     3  LGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLS   62
              :||::|||||||||||||||||||||||||||||||||||||||||||||
Db    39  LPPSLPLPRSFLLKSLEQVRKIQASGSVLLEQLCATYKLCHPEELVLLGHSLGIPKASLS   98

Qy    63  SCPSQALAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQME   122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    99  GCSSQAIQQTQCLSQLHSGLCLCYQGLLQALSGISPALAPTLDLLQLDVANFATTIWQQME  158

Qy   123  ELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLA   172
              ||||:|:|||||||||||||||||||||||||:||:||| ||||
Db   159  NLGVAPTVQPTQSAMPAFTSAFQRRAGGVLAISYLQGFLETARLALHHLA   208
```

```
RESULT 8
MGF_CHICK
ID   MGF_CHICK    STANDARD;    PRT;    201 AA.
AC   P13854;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Myelomonocytic growth factor precursor (MGF).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89231616; PubMed=2785450;
RA   Leutz A., Damm K., Sterneck E., Kowenz E., Ness S., Frank R.,
RA   Gausepohl H., Pan Y.-C.E., Smart J., Hayman M., Graf T.;
RT   "Molecular cloning of the chicken myelomonocytic growth factor (cMGF)
RT   reveals relationship to interleukin 6 and granulocyte colony
RT   stimulating factor.";
RL   EMBO J. 8:175-181(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92195319; PubMed=1549124;
RA   Sterneck E., Blattner C., Graf T., Leutz A.;
RT   "Structure of the chicken myelomonocytic growth factor gene and
RT   specific activation of its promoter in avian myelomonocytic cells by
RT   protein kinases.";
RL   Mol. Cell. Biol. 12:1728-1735(1992)
CC   -!- FUNCTION: HEMATOPOIETIC GROWTH FACTOR THAT STIMULATES THE
CC       PROLIFERATION AND COLONY FORMATION OF NORMAL AND TRANSFORMED AVIAN
CC       CELLS OF THE MYELOID LINEAGE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; M85034; AAA48694.1; -.
DR   EMBL; X14477; CAA32639.1; -.
DR   PIR; A42247; A42247.
DR   HSSP; P35834; 1BGD.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
```

```
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Growth factor; Glycoprotein; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    201            MYELOMONOCYTIC GROWTH FACTOR.
FT   DISULFID     61     67            BY SIMILARITY.
FT   DISULFID     89     99            BY SIMILARITY.
FT   CARBOHYD    123    123            N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    137    137            N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE    201 AA;  22373 MW;  240A8DD21B4244E6 CRC64;

Query Match             33.9%;  Score 304;  DB 1;  Length 201;
Best Local Similarity   40.6%;  Pred. No. 6.6e-21;
Matches 67;  Conservative 28;  Mismatches 68;  Indels 2;  Gaps 1;

Qy    11  QSFLLKLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQ   70
              |||||:|||||||||||||||||||||||||||||||||||||||||||||
Db    36  QLFLHKNLEFTRKIRGDVAALQRAVCDTFQLCTEEELQLVQPDPHLVQAPLDQCHKRGFQ   95

Qy    71  LAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEBLGMAPAL  130
              :|: ||||||||:||||||||||||: || |||||:|| ||||||||||||
Db    96  AEVCFTQIRAGLHAYHDSLGAVLRLLPNHTTLVETIQLDAANLSSNIQQQMEDLGLDTVT  155

Qy   131  QPTQ--GAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQ   173
              |:|  :| :|: || ||| : |||| | || :
Db   156  LPAEQRSPPPTFSGPFQQQVGGFFILANFQRFLETAYRALRHLAR   200
```

```
RESULT 9
IL6_PIG
ID   IL6_PIG    STANDARD;    PRT;    212 AA.
AC   P26893; Q95KN6;
DT   01-AUG-1992 (Rel. 23, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Interleukin-6 precursor (IL-6).
GN   IL6.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91338547; PubMed=1873476;
RA   Richards C., Saklatvala J.;
RT   "Molecular cloning and sequence of porcine interleukin 6 cDNA and
RT   expression of mRNA in synovial fibroblasts in vitro.";
RL   Cytokine 3:269-276(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92360284; PubMed=1497880;
RA   Mathialagan N., Bixby J.A., Roberts M.R.;
RT   "Expression of interleukin-6 in porcine, ovine, and bovine
RT   preimplantation conceptuses.";
RL   Mol. Reprod. Dev. 32:324-330(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Landrace x Meishan; TISSUB=Blood;
RA   Liu S., Meng M., Gao R.;
RT   "Cloning and expression of interleukin 6 gene from Landrance x Meishan
RT   hybrid swine.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: IL-6 is a cytokine with a wide variety of biological
CC       functions; it plays an essential role in the final differentiation
CC       of B-cells into Ig-secreting cells, it induces myeloma and
CC       plasmacytoma growth, it induces nerve cells differentiation, in
CC       hepatocytes it induces acute phase reactants.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M86722; AAC37333.1; -.
DR   EMBL; M80258; AAC27127.1; -.
DR   EMBL; AF309651; AAG27730.1; -.
DR   PIR; I46590; I46590.
DR   PIR; I46621; I46621.
DR   HSSP; P05231; 1ALU.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Acute phase; Cytokine; Growth factor; Glycoprotein; Signal.
FT   SIGNAL        1     29       BY SIMILARITY.
FT   CHAIN        30    212       INTERLEUKIN-6.
FT   DISULFID     72     78       BY SIMILARITY.
FT   DISULFID    101    111       BY SIMILARITY.
FT   CONFLICT     30     30       E -> G (IN REF. 1).
SQ   SEQUENCE   212 AA;   23952 MW;  1E736FB230B4FC5D CRC64;

Query Match             11.9%;  Score 106.5;  DB 1;  Length 212;
  Best Local Similarity   22.6%;  Pred. No. 0.0098;  -
  Matches   36;  Conservative   33;  Mismatches   85;  Indels    5;  Gaps    3;

Qy   16 KCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP-WAPLSSCPSQALQLAGC 74
         | :|:||||::::|:|  |:|
Db   52 KTEELIKYILGKISAMRKEMCEXYEKCENSKEVLAENNLNLPKVAEKDGCFQSGFNQETC 111

Qy   75 LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPT- 133
         | :: :|||  |||  :|: :|:  ||
Db  112 LMRITTGLVEFQIYLDYILQKEYESNKGNVEAVQISTKALIQTLRQGKNPDKATTPNPTT 171

Qy  134 -QGAMPAFAS--AFQRRAGGVLVASHLQSFLEVSYRVLR 169
          :|:: |  |:|||||
Db  172 NAGLLDKLQSQNEWMKNTKIILILRSLEDFLQFSLRAIR 210

RESULT 10
IL6_ORCOR
ID   IL6_ORCOR      STANDARD;      PRT;   205 AA.
AC   Q28747;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Interleukin-6 precursor (IL-6) (Fragment).
GN   IL6.
OS   Orcinus orca (Killer whale).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC   Orcinus.
OX   NCBI_TaxID=9733;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96163018; PubMed=8575817;
RA   King D.P., Schrenzel M.D., McKnight M.L., Reidarson T.H., Hanni K.D.,
RA   Stott J.L., Ferrick D.A.;
RT   "Molecular cloning and sequencing of interleukin 6 cDNA fragments from
RT   the harbor seal (Phoca vitulina), killer whale (Orcinus orca), and
RT   Southern sea otter (Enhydra lutris nereis).";
RL   Immunogenetics 43:190-195(1996).
CC   -!- FUNCTION: IL-6 is a cytokine with a wide variety of biological
CC       functions; it plays an essential role in the final differentiation
CC       of B-cells into Ig-secreting cells, it induces myeloma and
CC       plasmacytoma growth, it induces nerve cells differentiation, in
CC       hepatocytes it induces acute phase reactants (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
```

```
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   -----------------------------------------------------------------------
DR   EMBL; L46803; AAB01429.1; -.
DR   HSSP; P05231; 1ALU.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Acute phase; Cytokine; Growth factor; Glycoprotein; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     21       BY SIMILARITY.
FT   CHAIN        22    205       INTERLEUKIN-6.
FT   DISULFID     64     70       BY SIMILARITY.
FT   DISULFID     93    103       BY SIMILARITY.
FT   CARBOHYD    164    164       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   205 AA;   23266 MW;  630BF3A457960832 CRC64;

Query Match             11.3%;  Score 101;  DB 1;  Length 205;
  Best Local Similarity   20.6%;  Pred. No. 0.03;  -
  Matches   32;  Conservative   34;  Mismatches   83;  Indels    6;  Gaps    2;

Qy   21 VRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP-WAPLSSCPSQALQLAGCLSQLH 79
         :|:||||::::|:|  |:|
Db   49 IKYILGKISAMRKEMCEKYDKCENSKEALAENNLNLPKMAEKDGCFQSGFNQETCLMRIT 108

Qy   80 SGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPTQGA--- 136
         |:: :|||  |||  :|: :|:  ||
Db  109 TGLLEYQ-YLDYLQNEYEGDKEAIEAVQISSKALAQILRQKVKNPDEVTTPDPTTNASLM 168

Qy  137 -MPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR 169
          :|:: |  |:|||||
Db  169 NNLQSQNDDWMKNTKIILILRSLENFLQFSLRAIR 203

RESULT 11
IL6_HORSE
ID   IL6_HORSE      STANDARD;      PRT;   208 AA.
AC   Q95I81; O19007; O46568;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Interleukin-6 precursor (IL-6).
GN   IL6.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20579380; PubMed=11137120;
RA   Swiderski C.E., Sobol G., Lunn D.P., Horohov D.W.;
RT   "Molecular cloning, sequencing, and expression of equine interleukin-
RT   6.";
RL   Vet. Immunol. Immunopathol. 77:213-220(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Leutenegger C.M., Huder J.B., von Rechenberg B., Akens M., Auer J.;
RT   "Molecular cloning of equine interleukin-6.";
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lai A.C.K.;
```

```
RT   "Cloning and expression of equine interleukin-6.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: IL-6 is a cytokine with a wide variety of biological
CC       functions: it plays an essential role in the final differentiation
CC       of B-cells into Ig-secreting cells, it induces myeloma and
CC       plasmacytoma growth, it induces nerve cells differentiation, in
CC       hepatocytes it induces acute phase reactants (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; U64794; AAB87703.1; -.
DR   EMBL; AF005227; AAB62246.1; -.
DR   EMBL; AF041975; AAC04574.1; -.
DR   PIR; T09216; T09216.
DR   HSSP; P05231; 1ALU.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Acute phase; Cytokine; Growth factor; Glycoprotein; Signal.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    208       INTERLEUKIN-6.
FT   DISULFID     69     75       BY SIMILARITY.
FT   DISULFID     98    108       BY SIMILARITY.
FT   CARBOHYD     71     71       N-LINKED (GLCNAC...).  (POTENTIAL).
FT   CARBOHYD    184    184       N-LINKED (GLCNAC...).  (POTENTIAL).
FT   CONFLICT      4      5       LS -> FF (IN REF. 1).
FT   CONFLICT      8      8       T -> A (IN REF. 2).
FT   CONFLICT    137    137       I -> V (IN REF. 3).
FT   CONFLICT    205    205       V -> I (IN REF. 3).
SQ   SEQUENCE   208 AA;  23325 MW;  A62F4C234056BF66 CRC64;

Query Match             11.2%;  Score 100;  DB 1;  Length 208;
Best Local Similarity   21.6%;  Pred. No. 0.038;
Matches   38;  Conservative   35;  Mismatches   95;  Indels    8;  Gaps    3;

Qy          2  PLGPASSLPQSFLL---KCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP  57
               ||| | : :  ||    ::    :    :|  | :::|
Db         31  PLGEDETTSNGPLLTTADKTKQHIKYILGKISALKNEMCNNFSKCENSKEVLAENNLNLP  90

Qy         58  -WAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATT  116
                ||:::  :::::|| :::|||:|| :  :|:::|
Db         91  KMAEKDGCFQSGFNQETCLMKITTGLSBFQIYLEYLQNEFKGEKENIKTMQISTKVLVQI  150

Qy        117  IWQQMEELGMA--PALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR  169
               | ||  || |   |:  ||  | :|| ::::|||
Db        151  LMQKMKNPEVTTPDPTAKSSLLAKLHSQNEWLKNTTTHLILRSLEDFLQFSLRAVR  206


RESULT 12
IL6_FELCA
ID   IL6_FELCA     STANDARD;      PRT;   208 AA.
AC   P41683;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Interleukin-6 precursor (IL-6).
GN   IL6.
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94162386; PubMed=8117820;
RA   Ohashi T., Matsumoto Y., Watari T., Goitsuka R., Tsujimoto H.,
RA   Hasegawa A.;
RT   "Molecular cloning of feline interleukin-6 cDNA.";
RL   J. Vet. Med. Sci. 55:941-944(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymphocytes;
RX   MEDLINE=94052249; PubMed=8234373;
RA   Bradley W.G., Gibbs C., Kraus L., Good R.A., Day N.K.;
RT   "Molecular cloning and characterization of a cDNA encoding feline
RT   interleukin-6.";
RL   Proc. Soc. Exp. Biol. Med. 204:301-305(1993).
CC   -!- FUNCTION: IL-6 is a cytokine with a wide variety of biological
CC       functions: it plays an essential role in the final differentiation
CC       of B-cells into Ig-secreting cells, it induces myeloma and
CC       plasmacytoma growth, it induces nerve cells differentiation, in
CC       hepatocytes it induces acute phase reactants.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; L16914; AAA16620.1; -.
DR   EMBL; D13227; BAA02507.1; -.
DR   PIR; I46084; I46084.
DR   HSSP; P05231; 1IL6.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Acute phase; Cytokine; Growth factor; Glycoprotein; Signal.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    208       INTERLEUKIN-6.
FT   DISULFID     68     74       BY SIMILARITY.
FT   DISULFID     97    107       BY SIMILARITY.
FT   CONFLICT      2      2       T -> N (IN REF. 2).
FT   CONFLICT     45     45       S -> P (IN REF. 2).
FT   CONFLICT    133    133       E -> K (IN REF. 2).
FT   CONFLICT    173    187       AKLQSQEEWLRHTTI -> LSCSHRRVAEAHNN (IN
FT                                REF. 2).
FT   CONFLICT    200    201       FS -> LR (IN REF. 2).
SQ   SEQUENCE   208 AA;  23401 MW;  93B4456B2989CA4C CRC64;

Query Match             10.6%;  Score 95;  DB 1;  Length 208;
Best Local Similarity   23.9%;  Pred. No. 0.11;
Matches   47;  Conservative   30;  Mismatches   70;  Indels   50;  Gaps    9;

Qy          2  PLG--PASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP-  57
               ||| :  :::  ::  ||||:  : :  :::|
Db         31  PLGGDATNRLPLTSADKMEELIKYILGKISALKKEMCDNYNKCEDSKEALAENNLNLPK  90

Qy         58  WAPLSSCPSQALQLAGCLSQLHSGL----------FL---YQG---------LLQAL  92
               ||:::  :::::|| :::|||:||:          :|    :|:
Db         91  LABKDGCFQSGFNQETCLRITTGLQEFQIYLKFLQDKYEGDEENAKSVYTSNVLLQML  150

Qy         93  EGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVL  152
               : :|   : |::
Db        151  K----RKGKNQDEVTIPV----PTV-----EVGLQAKLQSQE-----EWLRHTTIHL  189

Qy        153  VASHLQSFLEVSYRVLR  169
               :: ||:  |:|  :|
```

```
Db          190 TLRRLEDFLQFSLRAVR 206

RESULT 13
HMCU_DROME
ID   HMCU_DROME     STANDARD;     PRT;   2175 AA.
AC   P10180; Q9W3Q6;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Homeobox protein cut.
GN   CT OR CG11387.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88232956; PubMed=2887632;
RA   Blochlinger K., Bodmer R., Jack J., Jan L.Y., Jan Y.N.;
RT   "Primary structure and expression of a product from cut, a locus
RT   involved in specifying sensory organ identity in Drosophila.";
RL   Nature 333:629-635(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: Regulator of cell fate decisions in multiple lineages.
CC       Specifically, functions as a determination factor that specifies
CC       sensory organ identity in precursor cells. Probably also involved
CC       in cell type specification of Malpighian tubules. In absence of
CC       cut gene external sensory organs are transformed into chordotonal
CC       organs.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
```

```
CC   -!- TISSUE SPECIFICITY: Detected in many cells in the central nervous
CC       system, all external sensory organs, some peripheral neurons, and
CC       in the non-neural cells of the spiracles and the Malpighian
CC       tubules.
CC   -!- DEVELOPMENTAL STAGE: Cell-specific pattern of expression. Broadly
CC       expressed during embryonic development.
CC   -!- DOMAIN: Asn at position 47 of the homeobox may participate in
CC       regulating DNA-binding activity by promoting homo- and
CC       heterodimerization.
CC   -!- SIMILARITY: Belongs to the CUT homeobox family.
CC   -!- SIMILARITY: Contains 1 homeobox domain.
CC   -!- SIMILARITY: Contains 3 CUT domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X07985; CAA30794.1; -.
DR   EMBL; AE003441; AAF46264.2; -.
DR   PIR; S03170; S03170.
DR   TRANSFAC; T02004; -.
DR   FlyBase; FBgn0004198; ct.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0008587; P:wing margin morphogenesis; NAS.
DR   InterPro; IPR007108; Cut_homeo.
DR   InterPro; IPR003350; Hmoeo_CUT.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF02376; CUT; 3.
DR   Pfam; PF00046; homeobox, 1.
DR   ProDom; PD000010; Homeobox, 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Transcription regulation; Homeobox; DNA-binding;
KW   Developmental protein; Nuclear protein; Repeat; Coiled coil.
FT   DOMAIN       265    343       COILED COIL (POTENTIAL).
FT   DOMAIN       433    499       COILED COIL (POTENTIAL).
FT   DNA_BIND     798    964       CUT 1.
FT   DOMAIN      1056   1161       COILED COIL (POTENTIAL).
FT   DNA_BIND    1329   1417       CUT 2.
FT   DOMAIN      1463   1522       COILED COIL (POTENTIAL).
FT   DNA_BIND    1608   1695       CUT 3.
FT   DNA_BIND    1745   1804       HOMEOBOX.
FT   DOMAIN       194    .210      ALA/GLN-RICH.
FT   DOMAIN       235    243       ALA-RICH.
FT   DOMAIN       271    293       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN       384    428       ASN-RICH.
FT   DOMAIN       547    554       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN       574    584       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN       616    630       ALA-RICH.
FT   DOMAIN       665    699       HIS/GLN-RICH (OPA-REPEAT).
FT   DOMAIN      2004   2014       ALA-RICH.
FT   DOMAIN      2071   2077       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      2124   2136       ALA/PRO-RICH.
SQ   SEQUENCE   2175 AA;  233628 MW;  08BF80C4861BD0AB CRC64;

  Query Match             89.5%;  Score 89.5;  DB 1;  Length 2175;
  Best Local Similarity   23.8%;  Pred. No. 4.8;
  Matches   35;  Conservative   20;  Mismatches   61;  Indels   31;  Gaps  3;

Qy          27 DGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAGCLSQLHSGLFLYQ 86
               |:|:|:|||||:||:::||||.:|:::|||||:|||::::::::::||::||::|:.|||| 
Db        1404 DENAVHKLVASQYKIA-PEKLMRTGSYSGSPQMP--------------------------Q 1437

Qy          87 GLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPTQGAMPAFASAFQR 146
               :||::.|:::|||.||:|.|:::||||||:||:|||||::|::||.|||||::::::|:|| 
Db        1438 GLASKMQAASLPMQKMMSLKLQEPAQAQHMQQMQQAAAMSAAMQQQ-----VAQAQQQ 1492

Qy         147 RAGGVLVASHLQSFLEVSYRVLRHLAQ 173
```

Db   1493 AQQAQQAQQHLQQQAQQHLQQQQHLAQ 1519
                 || :   : ||||
Db   151 IQILKQKIADLITTPATNTDLLEKMQSSNEWVKNAKIILILRNLENFLQFSLRAIR 206

RESULT 14
ID   IL6_BOVIN   STANDARD;   PRT;   208 AA.
AC   P26892;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Interleukin-6 precursor (IL-6).
GN   IL6.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein;
RX   MEDLINE=93076003; PubMed=1446077;
RA   Droogmans L., Cludts I., Cleuter Y., Kettmann R., Burny A.;
RT   "Nucleotide sequence of bovine interleukin-6 cDNA.";
RL   DNA Seq. 2:411-413(1992).
CC   -!- FUNCTION: IL-6 is a cytokine with a wide variety of biological
CC   functions: it plays an essential role in the final differentiation
CC   of B-cells into Ig-secreting cells, it induces myeloma and
CC   plasmacytoma growth, it induces nerve cells differentiation, in
CC   hepatocytes it induces acute phase reactants.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X57317; CAA40572.1; -.
DR   PIR; A56610; A56610.
DR   HSSP; P05231; 1IL6.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Acute phase; Cytokine; Growth factor; Glycoprotein; Signal.
FT   SIGNAL   1   29   BY SIMILARITY.
FT   CHAIN   30   208   INTERLEUKIN-6.
FT   DISULFID   72   78   BY SIMILARITY.
FT   DISULFID   101   111   BY SIMILARITY.
FT   CARBOHYD   38   38   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   208 AA; 23758 MW; ADF000B9BA2EC341 CRC64;

Query Match   9.6%; Score 86; DB 1; Length 208;
Best Local Similarity 20.5%; Pred. No. 0.74;
Matches   36; Conservative 37; Mismatches 95; Indels   8; Gaps   3;

Qy   2 PLG--PASSLPQSFLLKCLEQ---VRKIQGDGAALQEKLCATYKLCHPELIVLLGHSL 54
             || | | ||| | :: ||||||||||||||||||||||||
Db   31 PLGEDFKNDTTPGRLLTTPEKTEALIKRMVDKISAMRKEICEKNDECSSKETLAENKL 90

Qy   55 GIP-WAPLSSCPSQALQAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTILQLDVADF 113
             ::: ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91 NLPKMEEKDGCFQSGFNQAICLIRTTAGLLEYQIYLDYLQNEYEGNQENVRDLRKNIRTL 150

Qy   114 ATTIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR 169
             :: :: :|: |||| : :|:||||||||||||||||||||||||||||||||

RESULT 15
ID   IL6_PHOVI   STANDARD;   PRT;   209 AA.
AC   Q28819;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Interleukin-6 precursor (IL-6) (Fragment).
GN   IL6.
OS   Phoca vitulina (Harbor seal).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Phoca.
OX   NCBI_TaxID=9720;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96163018; PubMed=8575817;
RA   King D.P., Schrenzel M.D., McKnight M.L., Reidarson T.H., Hanni K.D.,
RA   Stott J.L., Ferrick D.A.;
RT   "Molecular cloning and sequencing of interleukin 6 cDNA fragments from
RT   the harbor seal (Phoca vitulina), killer whale (Orcinus orcal), and
RT   Southern sea otter (Enhydra lutris nereis).";
RL   Immunogenetics 43:190-195(1996).
CC   -!- FUNCTION: IL-6 is a cytokine with a wide variety of biological
CC   functions: it plays an essential role in the final differentiation
CC   of B-cells into Ig-secreting cells, it induces myeloma and
CC   plasmacytoma growth, it induces nerve cells differentiation, in
CC   hepatocytes it induces acute phase reactants (BY similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; L46802; AAB01430.1; -.
DR   HSSP; P05231; 1IL6.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Acute phase; Cytokine; Growth factor; Glycoprotein; Signal.
FT   NON_TER   1   1
FT   SIGNAL   <1   26   BY SIMILARITY.
FT   CHAIN   27   209   INTERLEUKIN-6.
FT   DISULFID   69   75   BY SIMILARITY.
FT   DISULFID   98   108   BY SIMILARITY.
SQ   SEQUENCE   209 AA; 23483 MW; 7514492E43B48E9 CRC64;

Query Match   9.4%; Score 84.5; DB 1; Length 209;
Best Local Similarity 22.0%; Pred. No. 1;
Matches   35; Conservative 30; Mismatches 89; Indels   5; Gaps   2;

Qy   16 KCLEQVRKIQGDGAALQEKLCATYKLCHPELIVLLGHSLGIP-WAPLSSCPSQALQLAGC 74
              :|:| |||:||||||||||||||||||||||||||||| :|: :
Db   49 KMEEFIKYILGKISALRKEMCDKYNKCEDSKEALAENLRLPKLAEKDGCFQSGFNQETC 108

Qy   75 LSQLHSGLFLYQGLLQAL---EGISPELGPTLDTILQLDVADFATTIWQQMEELGMAPAL 130
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   109 LTRITTGLLEFQIHLKYIQANYEGNKEDANSVYISTKLLVQMLMKKVKSQDEVTTPDPTT 168

Qy   131 QPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR 169
              :| :||||||||||||||||||||||||||||||||||
Db   169 DTSLQAILKAQDKWLKHTTIHLILRSLEDFIQFSLRAVR 207

us-10-009-792c-19.rsp

Tue Oct 12 09:18:49 2004

Search completed: October 6, 2004, 16:04:29
Job time : 25 secs

This Page Blank (uspto)

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:   October 6, 2004, 15:58:27 ; Search time 117 Seconds
                                      (without alignments)
                             469.232 Million cell updates/sec

Title:           US-10-009-792C-19
Perfect score:   896
Sequence:        1 TPLGPASSLPQSFLLKCLEQ........SHLQSFLEVSYRVLRHLAQP 174

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:    1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_25:*
                 1: sp_archea:*
                 2: sp_bacteria:*
                 3: sp_fungi:*
                 4: sp_human:*
                 5: sp_invertebrate:*
                 6: sp_mammal:*
                 7: sp_mhc:*
                 8: sp_organelle:*
                 9: sp_phage:*
                10: sp_plant:*
                11: sp_rodent:*
                12: sp_virus:*
                13: sp_vertebrate:*
                14: sp_unclassified:*
                15: sp_rvirus:*
                16: sp_bacteriap:*
                17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	896	100.0	200	4	Q8N4W3	Q8n4w3 homo sapien
2	744	83.0	195	6	Q9GGU0	Q9gju0 felis silve
3	634	70.8	214	11	P97712	P97712 rattus norv
4	511	57.0	127	6	Q8MKE0	Q8mke0 equus cabal
5	111	12.4	241	13	Q90YI0	Q90yi0 gallus gall
6	110.5	12.3	212	6	Q8MJ75	Q8mj75 sus scrofa
7	108	12.1	208	6	Q9XT80	Q9xt80 delphinapte
8	96	10.7	189	4	Q9H2A5	Q9h2a5 homo sapien
9	94.5	10.5	455	16	Q8D706	Q8d706 vibrio vuln
10	94	10.5	345	16	Q9KTL6	Q9ktl6 vibrio chol
11	93	10.4	189	4	Q9NPF7	Q9npf7 homo sapien
12	89.5	10.0	214	6	Q8MKE5	Q8mke5 sus scrofa
13	88.5	9.9	666	16	Q9A523	Q9a523 caulobacter
14	88.5	9.9	786	16	Q9I019	Q9i019 pseudomonas
15	88.5	9.9	1931	2	Q8RJY3	Q8rjy3 stigmatella
16	86	9.6	502	8	Q34008	Q34008 beta vulgar
17		86	9.6	788	11	Q8CF87	Q8cf87 mus musculu
18		86	9.6	850	11	Q7TQ21	Q7tq21 mus musculu
19		86	9.6	851	11	Q8CF88	Q8cf88 mus musculu
20		86	9.6	852	11	Q811T9	Q811t9 mus musculu
21		85.5	9.5	211	6	Q865W7	Q865w7 camelus bac
22		85.5	9.5	1288	10	Q8LQR8	Q8lqr8 oryza sativ
23		85	9.5	193	6	Q9N2H9	Q9n2h9 sus scrofa
24		84.5	9.4	211	6	Q865X6	Q865x6 lama glama
25		84	9.4	175	6	Q9TTH4	Q9tth4 aotus nigri
26		84	9.4	209	6	Q9TTH3	Q9tth3 aotus lemur
27		84	9.4	274	17	O29363	O29363 archaeoglob
28		84	9.4	290	4	Q9P0S7	Q9p0s7 homo sapien
29		84	9.4	641	17	Q8ZTH1	Q8zth1 pyrobaculum
30		83.5	9.3	212	6	Q8MKH0	Q8mkh0 saimiri sci
31		83.5	9.3	846	11	Q923K8	Q923k8 cavia porce
32		83.5	9.3	846	11	Q8CHL9	Q8chl9 cavia porce
33		82.5	9.2	207	6	Q9MYZ7	Q9myz7 canis famil
34		82.5	9.2	207	6	Q28403	Q28403 enhydra lut
35		82.5	9.2	327	13	Q7ZTY0	Q7zty0 brachydanio
36		81	9.0	244	2	Q8GMH8	Q8gmh8 streptomyce
37		81	9.0	5457	2	Q83VS1	Q83vs1 pseudomonas
38		80.5	9.0	641	4	O94927	O94927 homo sapien
39		80.5	9.0	3729	2	O33956	O33956 streptomyce
40		80	8.9	355	10	Q9AV51	Q9av51 oryza sativ
41		80	8.9	355	10	Q7XDB0	Q7xdb0 oryza sativ
42		80	8.9	474	16	Q9RXX9	Q9rxx9 deinococcus
43		80	8.9	477	4	Q9NT11	Q9nt11 homo sapien
44		80	8.9	811	16	Q8DG52	Q8dg52 synechococc
45		80	8.9	966	4	Q8NDG1	Q8ndg1 homo sapien

ALIGNMENTS

RESULT 1
Q8N4W3
ID   Q8N4W3   PRELIMINARY;   PRT;   200 AA.
AC   Q8N4W3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC033245; AAH33245.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005125; F:cytokine activity; IEA.
DR   GO; GO:0008083; F:growth factor activity; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   200 AA; 21543 MW; 8648AA5SB329A96C CRC64;

Query Match   100.0%; Score 896; DB 4; Length 200;
Best Local Similarity   100.0%; Pred. No. 1.9e-77;
Matches 174; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   27 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 86
```

```
Qy       61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       87 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 146

Qy      121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      147 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 200
```

RESULT 2
```
ID   Q9GJU0;
AC   Q9GJU0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)            PRELIMINARY;       PRT;    195 AA.
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Granulocyte colony-stimulating factor precursor.
GN   G-CSF.
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamamoto A., Iwata A., Tuchiya K., Katsumata A., Oishi K.,
RA   Tsujimoto H., Hasegawa A., Ueda S.;
RT   "Molecular cloning and expression of the cDNA encoding feline
RT   granulocyte colony-stimulating factor.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yamamoto A., Iwata A., Tuchiya K., Katsumata A., Oishi K., Saito T.,
RA   Tsujimoto H., Hasegawa A., Ueda S.;
RT   "Molecular cloning and expression of the cDNA encoding feline
RT   granulocyte colony-stimulating factor.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB042552; BAB17789.1; -.
DR   EMBL; AB042553; BAB17757.1; -.
DR   HSSP; P35834; 1BGE.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005125; F:cytokine activity; IEA.
DR   GO; GO:0008083; F:growth factor activity; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Signal.
FT   SIGNAL         1     21       POTENTIAL.
SQ   SEQUENCE   195 AA;  21255 MW;  544C682909412FCF CRC64;

Query Match             83.0%;  Score 744;  DB 6;  Length 195;
Best Local Similarity   81.0%;  Pred. No. 5.7e-63;  -;
Matches  141;  Conservative  11;  Mismatches  22;  Indels  0;  Gaps  0;

Qy        1 TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAP 60
            |||||||:|||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       22 TPLGPTSSLPQSFLLKCLEQVRKVQADGTALQERLCAAHKLCHPEELVLLGHALGIPQAP 81

Qy       61 LSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQ 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db       82 LSSCSSQALQLTGCLRQLHSGLFLYQGLLQALAGLSPELAPTLDMLQLDITDFAINIWQQ 141

Qy      121 MEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
            |||||||||||||||||||||||||||||||||||::||||||||||||||||:
Db      142 MEDVGMAPAVPPTQGTMPTFTSAFQRRAGGTLVASNLQSFLEVAYRALRHFTKP 195
```

RESULT 3

```
ID   P97712;
AC   P97712;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)            PRELIMINARY;       PRT;    214 AA.
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Granulocyte colony stimulating factor.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97074656; PubMed=8917083;
RA   Han S.W., Ramesh N., Osborne W.R.;
RT   "Cloning and expression of the cDNA encoding rat granulocyte colony-
RT   stimulating factor.";
RL   Gene 175:101-104(1996).
DR   EMBL; U37101; AAC52915.1; -.
DR   PIR; JC5043; JC5043.
DR   HSSP; P09919; 1RHG.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005125; F:cytokine activity; IEA.
DR   GO; GO:0008083; F:growth factor activity; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR003629; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
SQ   SEQUENCE   214 AA;  23659 MW;  29B88B17B684C55 CRC64;

Query Match             70.8%;  Score 634;  DB 11;  Length 214;
Best Local Similarity   73.8%;  Pred. No. 1.9e-52;  -;
Matches  127;  Conservative  10;  Mismatches  35;  Indels  0;  Gaps  0;

Qy        3 LGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLS 62
            |:||:||||:||||||||||:||||||:|||||||||||||||||||||||||||||||:
Db       30 LPPSLPLPRSFLLKSLEQVRKIQARNTELLEQLCATYKLCHPEELVLFGHSLGIPKASLS 89

Qy       63 SCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQME 122
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       90 SCSSQALQQTKCLSQLHSGLFLYQGLLQALAGISSELAPTLDMLHLDVDNFATTIWQQME 149

Qy      123 ELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP 174
            ||||||||:||||||||||||||||||||||:: ||||||||||||||||| :
Db      150 SLGVAPTVQPTQSTMPIFTSAFQRRAGGVLVTSYLQSFLETAHHALHHLPRP 201
```

RESULT 4

```
ID   Q8MKE0;
AC   Q8MKE0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)            PRELIMINARY;       PRT;    127 AA.
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Granulocyte colony-stimulating factor.
GN   G-CSF.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Noronha L.E., Takafuji V.A., Sharova L.V., Crisman M.V., Howard R.D.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF503365; AAM34205.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005125; F:cytokine activity; IEA.
DR   GO; GO:0008083; F:growth factor activity; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
```

```
DR   InterPro; IPR003529; GCSF_MGF.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD008388; GCSF_MGF; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
SQ   SEQUENCE 127 AA; 13657 MW; 25559C256980202077 CRC64;

Query Match             57.0%;  Score 511; DB 6; Length 127;
Best Local Similarity   79.5%;  Pred. No. 5.5e-41;
Matches 101; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

Qy  48  VLLGHSLGIPWAPLSSCPSQALQAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQ 107
            |||||||||  ||||||:||||| :|:|||||||||:|||||||||||||||| ||||||
Db   1  MLLGHSLGIPQPLLSSCSSQALQLTGCLSQLHSGLLLYQGLLQALEGISPELAPTLDMLQ  60

Qy 108  LDVADFATTIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRV 167
            |||:||||.|||||| |||| ||||  |.||.||||||||||||||||:||.||| :||.
Db  61  LDVTDFATNIWQQMEDLGVAPVQPTHGPMPTFASAFQRRAGGVLVASNLQRFLELAYRG 120

Qy 168  LRHLAQP 174
            ||:|:|
Db 121  LRYLAEP 127
```

RESULT 5
Q90YI0

```
ID   Q90YI0      PRELIMINARY;      PRT;   241 AA.
AC   Q90YI0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Interleukin-6 precursor.
GN   IL-6.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schneider K., Klaas R., Kaspers B., Staeheli P.;
RT   "chicken interleukin-6: cDNA structure and biological properties.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kaiser P., Rothwell L., Galyov E.E., Barrow P.A., Burnside J.,
RA   Wigley P.;
RT   "Differential cytokine expression in avian cells in response to
RT   invasion by Salmonella typhimurium, Salmonella enteritidis and
RT   Salmonella gallinarum.";
RL   Microbiology 146:3217-3226(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Kaiser P.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ309540; CAC40812.1; -.
DR   EMBL; AJ250838; CAC15566.2; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005125; F:cytokine activity; IEA.
DR   GO; GO:0005138; F:interleukin-6 receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
KW   Signal.
FT   SIGNAL     1     47       POTENTIAL.
FT   CHAIN     48    241       MATURE CHIL-6.
SQ   SEQUENCE 241 AA; 26790 MW; 65778049F25BD2F8 CRC64;

Query Match             12.4%;  Score 111; DB 13; Length 241;
Best Local Similarity   22.2%;  Pred. No. 0.015;
Matches 46; Conservative 35; Mismatches 74; Indels 52; Gaps 7;

Qy   3  LGPASSLP---------------QSFLLKCLEQVRKIQGDGAALQEKLCATYKLCH  43
            |:|::|
Db  41  LPPAAAVPLPAAADSSGEVGLEEEAGARRAILDCEPLARVLRDRAVQLQDEMCKKFTVCE 100

Qy  44  PEELVLLGHSLGIPWAPLSSCPSQALQLAG-----CLSQLHSGLFLYQGLLQALEGISPE  98
            ||:||::  :::|::|              |||||||||||||||||||||||||
Db 101  NSMEMLVRNNLNLP----KVTEEDGCLLAGFDEEKCLTKLSSGLFAFQTYLEFIQETFDS 156

Qy  99  LGPTLDTLQLDVADFATTIWQQMEELGMAPALQPTQGAMPAFASAFQRRAGGVL------ 152
            |||::|||||||||||||||||||||||||||||||||||||||||||
Db 157  EKQNVESLCYSTKHLAATIRQMV--------INPDEVVIP--DSAAQKSLLANLKSDKDW 206

Qy 153  ---VASHL----QSFLEVSYRVLRHL 171
                        |:||||||||:|
Db 207  IEKITMHLILRDFTSFMEKIVRAVRYL 233
```

RESULT 6
Q8MJ75

```
ID   Q8MJ75      PRELIMINARY;      PRT;   212 AA.
AC   Q8MJ75;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   IL-6.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lee D., Yoo H., Choi I.;
RT   "Sequencing of Porcine IL-6 in Korea.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF518322; AAM74938.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005125; F:cytokine activity; IEA.
DR   GO; GO:0005138; F:interleukin-6 receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
SQ   SEQUENCE 212 AA; 23881 MW; 1F540E7030BCFD77 CRC64;

Query Match             12.3%;  Score 110.5; DB 6; Length 212;
Best Local Similarity   23.3%;  Pred. No. 0.015;
Matches 37; Conservative 32; Mismatches 85; Indels 5; Gaps 3;

Qy  16  KCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP-WAPLSSCPSQALQLAGC  74
            |::|     |::|| ::|                |:|| |||||||::|:
Db  52  KTEELIKYILGKISAMREMCEKYEKCENSKEVLAENNLNLPKMAEKDGCFQSGFNQETC 111

Qy  75  LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPT- 133
            |::|||  |:: |::|   |::| ::|::|:|:
Db 112  LMRITTGLVEFQIYLDYLQKEYESNKGNVEAVQISTKALIQTLRQKGKNPDKATTPNFTT 171

Qy 134  -QGAMPAFAS-AFQRRAGGVLVASHLQSFLEVSYRVLR 169
              :|  ||:    |||||||||:|::
Db 172  NAGLDKLQSQNEWMKNTKILILRSLEDFIQFSLRAIR 210
```

RESULT 7
Q9XT80

```
ID   Q9XT80          PRELIMINARY;   PRT;   208 AA.
AC   Q9XT80;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Interleukin 6.
OS   Delphinapterus leucas (Beluga whale).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
OC   Monodontidae; Delphinapterus.
OX   NCBI_TaxID=9749;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   St-Laurent G., De Guise S., Fournier M., Archambault D.;
RT   "Molecular cloning and phylogenetic analysis of beluga whale
RT   (Delphinapterus leucas) interleukin 6.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF076643; AAD42929.1; -.
DR   HSSP; P05231; 1ALU.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005125; F:cytokine activity; IEA.
DR   GO; GO:0005138; F:interleukin-6 receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
SQ   SEQUENCE  208 AA;  23456 MW;  81CC85C6E80389C4 CRC64;

Query Match             12.1%;  Score 108; DB 6; Length 208;
Best Local Similarity   21.3%;  Pred. No. 0.025;
Matches   35;  Conservative 35; Mismatches  81; Indels  6; Gaps  2;

Qy       21 VRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP-WAPLSSCPSQALQLAGCLSQLH 79
            :  :|||   |||||:| |   |||| :|||   :    | ||||||||||  |:|| ||
Db       52 IKYILGKISAMRKEMCEKYDKCENSKEALAENNLNLPKMAEKDGCFQSGFNQETCLMRIT 111

Qy       80 SGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPTQGA-- 136
            |||||||| : ||||:|| :|| |||  :   |||| :|  ||  ||:| |||| ||
Db      112 TGLLEYQIYLDYILQNEYEGDKGSIRAVQISIKALAQILRQKVKQNPDEVTTPDTTNASLM 171

Qy      137 --MPAFASAFQRRAGGVLVASHLQSFLEVSYRVLR 169
              :|||||||||||||| :|| |:|  | || |
Db      172 NNLQSQNDDWMRNTKILILRSLENFLQFSLRAVR 206


RESULT 8
Q9H2A5
ID   Q9H2A5          PRELIMINARY;   PRT;   189 AA.
AC   Q9H2A5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Interleukin 23 p19 subunit.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20567322; PubMed=11114383;
RA   Oppmann B., Lesley R., Blom B., Timans J.C., Xu Y., Hunte B., Vega F.,
RA   Yu N., Wang J., Singh K., Zonin F., Vaisberg E., Churakova T.,
RA   Liu M.-R., Gorman D., Wagner J., Zurawski S., Liu Y.-J., Abrams J.S.,
RA   Moore K.W., Rennick D., de Waal-Malefyt R., Hannum C., Bazan J.F.,
RA   Kastelein R.A.;
RT   "Novel p19 Protein Engages IL-12p40 to Form a Cytokine, IL-23, with
RT   Biological Activities Similar as Well as Distinct from IL-12.";
RL   Immunity 13;715-725(2000).
DR   EMBL; AF301620; AAG37232.1; -.
DR   Genew; HGNC:15488; IL23A.
SQ   SEQUENCE  189 AA;  20744 MW;  BFB5C0F42D4C1E3A CRC64;

Query Match             10.7%;  Score 96; DB 4; Length 189;
Best Local Similarity   26.1%;  Pred. No. 0.31;
Matches   48;  Conservative 19; Mismatches  77; Indels  40; Gaps  9;

Qy       14 LLKCLEQVRKIQGDGA------ALQEKLCATYKLCHPEELVLLGH----------SLG 55
            |:  | :    | | |      |||:| ||||||||||||||||          |||
Db       12 LLPWTAQGRAVPGGSSPAWTQCCQQLSQKLCTLAWSAHP---LVGHMDLREBGDEETND 67

Qy       56 IPWAPL-SSCPSQALQ-LAGCLSQLHSGLFLYQGLL--QALEGISPELGPTLDTLQLDV 110
            |  | | |||||:||| |||| |:|||||||||| |  ||||||||||||||||| |||
Db       68 VPHIQCGCDPGQGLRDNSQFCLQRIHQGLIFYEKLLGSDIFTG-EPSLLPDSPVAQLHA 126

Qy      111 ADFATTIWQQME---ELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYR 166
            ||||||||||:    |||||||||  ||:  ||||||||||||||||||||||||||:|
Db      127 SLLGLSQLLQPEGHHWETQQIPSLSPSQ---------PWQRLLLRFKILRSLQAFVAVAAR 178

Qy      167 VLRH 170
            | |
Db      179 VFAH 182


RESULT 9
Q8D706
ID   Q8D706          PRELIMINARY;   PRT;   455 AA.
AC   Q8D706;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Biopolymer transport protein.
GN   VV20363.
OS   Vibrio vulnificus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=672;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CMCP6;
RA   Rhee J.H., Kim S.Y., Kim J.J., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA   Choy H.E.;
RT   "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE016809; AAO07321.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008565; F:protein transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR002898; MotA_ExbB.
DR   Pfam; PF01618; MotA_ExbB; 1.
KW   Complete proteome.
SQ   SEQUENCE  455 AA;  49446 MW;  241A87B1CDB8B942 CRC64;

Query Match             10.5%;  Score 94.5; DB 16; Length 455;
Best Local Similarity   24.6%;  Pred. No. 1.2;
Matches   43;  Conservative 15; Mismatches  40; Indels  77; Gaps  7;

Qy        6 ASSLPQSFLLKCLEQVRKIQ---GDGAALQEKLCATYKLCHPEELVLLGH--------52
            |  | |  |   | :|| |    | ::| ::||||||||||||||||||
Db      140 AQSLPDIAQLKTLWHVMQQQIVASGEMRALQVRLINAEGVVEPTDVMALGNFALLAPQGF 199

Qy       53 ------SLGIPWAPLSSCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDT 105
                  ||| |||||  | ||||  | |:||| |||| | |  | ||||:|| |||||
Db      200 VNWNSDKGFAIPYAKL------------------------------------PEGAPTLDT 224

Qy      106 LQLDVADFATTIWQQMEELGMAPAL-QPTQGAM-------PAFASAFQRRAGGVL 152
            | |:|  ||| |  ||   |:|| | |||| |         ||||| |||||||
Db      225 L--------RSQHEISMTPVLIDPTRGVLLEQYAHSPSLADRF--AAGGVV 265


RESULT 10
Q9KTL6
ID   Q9KTL6          PRELIMINARY;   PRT;   345 AA.
```

```
AC   Q9XTL6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein VC0886.
GN   VC0886.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=El Tor N16961 / Serotype O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
DR   EMBL; AE004172; AAF94048.1; -.
DR   PIR; C82270; C82270.
DR   TIGR; VC0886; -.
DR   InterPro; IPR001303; Aldolase_II_N.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE 345 AA; 38781 MW; E446D7798ADA7E01 CRC64;

Query Match            10.5%;  Score 94; DB 16;  Length 345;
Best Local Similarity  32.2%;  Pred. No. 0.97;
Matches   38;  Conservative  15;  Mismatches  43;  Indels  22;  Gaps   6;

Qy      7 SSLPQSFLLKCLEQVRKIQGDGAALQEKLC----ATYKLCHPEELVLLGHSLGIPWAPLS 62
            ::|||||||||||  :||        |||||||  :||:||:|
Db    228 ASDPDLFLLAAL--VRALAGAPANILDATCDAILASPAICHPEVLIALA--GRSWMSLE 282

Qy     63 S---CPSQALQLAGCLSQ-LHSGLFLYQGLLQALEGI-------SPELGPTLDTLQ 107
              :  || |:|| |::   |:|  | | ::||         | ||
Db    283 NSGRAQRFLIRLAQTQNQPLFNQLFADVVMLPALRGVMLPLLHASPSPELACALEQLQ 340


RESULT 11
Q9NPF7
ID   Q9NPF7       PRELIMINARY;       PRT;   189 AA.
AC   Q9NPF7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   SGRF precursor.
GN   SGRF;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RA   Hirata Y., Kosuge Y.;
RT   "SGRF, a novel member of the IL-6/G-CSF family.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB030000; BAA93686.1; -.
DR   EMBL; AB030001; BAA93687.1; -.
KW   Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    189       SGRF.
SQ   SEQUENCE 189 AA; 20730 MW; 51B5C0F188EC1B9F CRC64;

Query Match            10.4%;  Score 93; DB 4;  Length 189;
Best Local Similarity  26.1%;  Pred. No. 0.59;
Matches   48;  Conservative  19;  Mismatches  77;  Indels  40;  Gaps   9;
```

```
Qy     14 LLKCLEQVRKIQGDGA------ALQEKLCATYKLCHPEELVLLGH------ 55
            |:||||||||||| :              :|||  ||:|||||||||
Db     12 LLPWTAQGRAVPGGSSPAWTCQCQLSQKLCTLAWSAHP----LVGHMDLREEGDEETND 67

Qy     56 IPWAPL-SSCPSQALQ--LAGCLSQLHSGLFLYQGLL--QALEGISPELGPTLDTLQLDV 110
            |  |||   :| |||  ||||  |:|||||:::|   |||:||:|||||||||:||||:
Db     68 VPHIQCGCDPGLRDNSQFCLQRIHQGLIFYEKLLGSDIFTG-EPSLLPDSPVGQLHA 126

Qy    111 ADFATTIWQQME---ELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYR 166
              |  :::|||          |:|: |       ||  |  |||:|  |:|||::|
Db    127 SLLGLSQLLQPEGHHWETQQIPSLSPSQ--------PWQRLLRFKILRSLQAFVAVAAR 178

Qy    167 VLRH 170
            |:|
Db    179 VFAH 182


RESULT 12
Q8MKE5
ID   Q8MKE5       PRELIMINARY;       PRT;   214 AA.
AC   Q8MKE5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Interleukin-6.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RA   Wu M., Gao R., Li J., Meng M., Long Z., Tang M., Liu S.;
RT   "Cloning and sequence analysis of interleukin-6 gene from Chenghua
RT   swine.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF493992; AAM27192.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005125; F:cytokine activity; IEA.
DR   GO; GO:0005138; F:interleukin-6 receptor binding; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR003573; IL6_MGF_GCSF.
DR   InterPro; IPR003574; Interleukin_6.
DR   Pfam; PF00489; IL6; 1.
DR   PRINTS; PR00433; IL6GCSFMGF.
DR   ProDom; PD004356; Interleukin_6; 1.
DR   SMART; SM00126; IL6; 1.
DR   PROSITE; PS00254; INTERLEUKIN_6; 1.
SQ   SEQUENCE 214 AA; 23765 MW; 50849FB04D0BDD7F CRC64;

Query Match            10.0%;  Score 89.5; DB 6;  Length 214;
Best Local Similarity  23.0%;  Pred. No. 1.5;
Matches   28;  Conservative  25;  Mismatches  68;  Indels   1;  Gaps   1;

Qy     16 KCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIP-WAPLSSCPSQALQLAGC 74
            | ||: |:|       : | | : | |||:||||  :|:|  |||:|| |:||:|
Db     52 KTELIKYILGKISAMRKEMCEKYEKCENSKEVLAENNLNLPKMAEKDGCFQSGFNQETC 111

Qy     75 LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPALQPTQ 134
            |||||||||| |||:|:||||||||||||||:||||    :||||||:||||||:|
Db    112 LMRITTGLVEFQIYLDYLQKEYBSNKGNVEAVQISTKALIQTLRQKGKNPDKATTPNPTT 171

Qy    135 GA 136
            |
Db    172 NA 173


RESULT 13
Q9A523
ID   Q9A523       PRELIMINARY;       PRT;   666 AA.
AC   Q9A523;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
```

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Oligopeptide transporter, OPT family.
GN   CC2646.
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA   Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA   Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR   EMBL; AE005932; AAK24613.1; -.
DR   PIR; A87577; A87577.
DR   TIGR; CC2646; -.
DR   InterPro; IPR004814; OPT fam.
DR   InterPro; IPR004813; Tetrpept_transpt.
DR   Pfam; PF03169; OPT; 1.
DR   TIGRFAMs; TIGR00728; OPT sfam; 1.
DR   TIGRFAMs; TIGR00733; TIGR00733; 1.
KW   Complete proteome.
SQ   SEQUENCE   666 AA;  67569 MW;  1ED3FF286CA8B6C0 CRC64;

Query Match             9.9%;  Score 88.5;  DB 16;  Length 666;
Best Local Similarity  32.5%;  Pred. No. 7.1;
Matches  37;  Conservative  14;  Mismatches  48;  Indels  15;  Gaps  4;

Qy   50 LGHSLGIPWAPLSSCPSQAIQLAGCLSQLHSGLFIYQGLIQALEGISPELGPTLDTLQLD 109

Db  196 LGAASGATGIGASS--SLAIMGAGHIMGITVGVAMPTGLFIAWAILVPILTLVTPMPEAD 253

Qy  110 VADFATTIWQ-QMEELG-----------MAPALQP-TQGAMPAFASAFQRRAGG 150
           |:|:||  :|||           ::|| |:  ||| :|| ||  ::|:
Db  254 AATHALITWKSQVRFLGAGVIGAAAIWTLAKLVGPITSGLKSAFAAAQARKAGG 307

RESULT 14
Q9IO19
ID   Q9IO19                  PRELIMINARY;     PRT;     786 AA.
AC   Q9IO19;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Probable sensor/response regulator hybrid.
GN   PA2824.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC       REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC   -!- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR   EMBL; AE004709; AAG06212.1; -.
DR   PIR; F83292; F83292.
DR   HSSP; P06143; 1AB6.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0000156; F:two-component response regulator activity; IEA.
DR   GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR   GO; GO:0007600; P:sensory perception; IEA.
DR   GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR   InterPro; IPR003594; ATPbind ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR005467; His_kinase.
DR   InterPro; IPR003661; His_kinA_N.
DR   InterPro; IPR001789; Response_reg.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00512; HisKA; 1.
DR   Pfam; PF00072; response_reg; 1.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   ProDom; PD000039; Response_reg; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00388; HisKA; 1.
DR   SMART; SM00448; REC; 1.
DR   PROSITE; PS50109; HIS_KIN; 1.
DR   PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW   Kinase; Phosphorylation; Sensory transduction; Transferase;
KW   Complete proteome.
SQ   SEQUENCE   786 AA;  85860 MW;  3B134513F3538F7C CRC64;

Query Match             9.9%;  Score 88.5;  DB 16;  Length 786;
Best Local Similarity  29.2%;  Pred. No. 8.6;
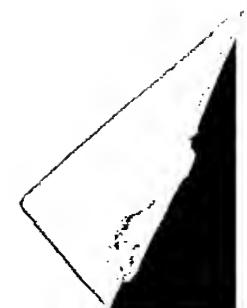Matches  50;  Conservative  19;  Mismatches  49;  Indels  53;  Gaps  10;

Qy   26 GDGAALQEKLC-ATYKLCHPEELVLLG--HSLGIPWAPLSSCPSQAIQLAG-CLSQLHSG 81

Db  490 GLGLALTRKLCEAMQGELTVESTVGLGSLFSVGLPLAPVSP-PLQALPLRGRVIAQCSAN 548

Qy   82 LFLYQGLLQALEGISPE------------LGPTLDTLQLDVDADFATTIWQQMEELGMA 127
                                    |||||||||:::: |:||  ||  || ||
Db  549 ----SGLAQLLQTWLPRWGLEYKLETDDSLLGHSLDVLISDCPDCL------MGLR 595

Qy  128 PALQPTQGAMPAFASAFQRRAGGVIVASHLQSFL-EVSYRV--LRHLAQP 174
           |:|:   |||  :||::: |:||| |
Db  596 PSI------------GTPILLVTAYGSFLPELARRLSPLRQLARP 629

RESULT 15
Q8RJY3
ID   Q8RJY3                  PRELIMINARY;     PRT;    1931 AA.
AC   Q8RJY3;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   StiD protein.
GN   STID.
OS   Stigmatella aurantiaca.
OC   Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC   Cystobacterineae; Cystobacteraceae; Stigmatella.
OX   NCBI_TaxID=41;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sg a15;
RA   Gaitatzis N., Silakowski B., Kunze B., Nordsiek G., Blocker H.,
RA   Hofle G., Muller R.;
RT   "The biosynthesis of the aromatic myxobacterial electron transport
RT   inhibitor stigmatellin is directed by a novel type of modular
RT   polyketide synthase.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ421825; CAD19088.1; -.
DR   GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
```

This Page Blank (uspto)

us-10-009-792c-19.rspt

```
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR001227; Ac_trans.
DR   InterPro; IPR000794; Ketoacyl_synth.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR006163; Pp_bind.
DR   InterPro; IPR000051; SAM_bind.
DR   InterPro; IPR002155; Thiolase.
DR   Pfam; PF00698; Acyl_transf; 1.
DR   Pfam; PF00109; ketoacyl-synt; 1.
DR   Pfam; PF02801; ketoacyl-synt_C; 1.
DR   Pfam; PF00550; pp-binding; 1.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
DR   PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR   PROSITE; PS00098; THIOLASE_1; 1.
KW   Phosphopantetheine; Transferase.
SQ   SEQUENCE   1931 AA;  208260 MW;  6DEAA0F496A9A84F CRC64;
```

```
Query Match       9.9%;   Score 88.5;  DB 2;  Length 1931;
Best Local Similarity   25.3%;   Pred. No. 25;
Matches   41;  Conservative   24;  Mismatches   48;  Indels   49;  Gaps    6;

Qy     6 ASSLPQSFILKCLEQVRKIQGDGAALQEKLCATYK---LCHPEELVLLGHSLGIPWAPLS 62
           |||::|::|         ||::|  : |::|      ||
Db   734 ASNGPRSTVL--------SGDPAALEQVLTSLRRQEVFCRPVDVDVASHS--------- 775

Qy    63 SCPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQME 122
           |::||||::       |:  ||       || :|::::|
Db   776 ------PQMEPLIPELERGLAVLQPRAAAVPFYSTVLGDVLDGRRLDASYWARNL----- 824

Qy   123 ELGMAPALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVS 164
           :||      :|||    :  | :||||
Db   825 ---RSPVL-------FFQSLQR----LIAAGHELFIEVS 849
```

```
Search completed: October  6, 2004, 16:06:34
Job time : 121 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    October  9, 2004, 01:28:34 ; Search time 27.6529 Seconds
                                        (without alignments)
                                        3612.319 Million cell updates/sec

Title:          US-10-009-792C-26
Perfect score:  180
Sequence:       1 atgtttaagtttaaaagaa........tcctgctcaagtgcttagag 180

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:   682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:     1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued_Patents_NA:*
            1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	89.2	49.6	642	1	US-08-104-445-2	Sequence 2, Appli
2	89.2	49.6	1413	1	US-08-104-445-1	Sequence 1, Appli
3	79	43.9	1349	1	US-09-570-856B-2	Sequence 1, Appli
4	79	28.2	1207	1	US-08-575-964-2	Sequence 2, Appli
5	50.8	28.2	1207	2	US-09-963-500-2	Sequence 2, Appli
6	50.8	28.1	546	2	US-08-469-318-177	Sequence 177, App
7	50.6	28.1	546	3	US-08-469-318-178	Sequence 178, App
8	50.6	28.1	546	3	US-08-468-609A-177	Sequence 177, App
9	50.6	28.1	546	3	US-08-468-609A-178	Sequence 178, App
10	50.6	28.1	546	4	US-08-446-872A-177	Sequence 177, App
11	50.6	28.1	546	4	US-08-446-872A-178	Sequence 178, App
12	50.6	28.1	546	4	US-08-762-227A-177	Sequence 177, App
13	50.6	28.1	546	4	US-08-762-227A-178	Sequence 178, App
14	50.6	28.1	546	5	PCT-US95-01185-177	Sequence 177, App
15	50.6	28.1	546	5	PCT-US95-01185-178	Sequence 178, App
16	50.6	28.1	906	3	US-08-469-318-53	Sequence 53, Appl
17	50.6	28.1	906	3	US-08-468-609A-53	Sequence 53, Appl
18	50.6	28.1	906	4	US-08-446-872A-53	Sequence 53, Appl
19	50.6	28.1	906	4	US-08-762-227A-53	Sequence 53, Appl
20	50.6	28.1	906	5	PCT-US95-01185-53	Sequence 53, Appl
21	50.6	28.1	921	3	US-08-469-318-72	Sequence 72, Appl
22	50.6	28.1	921	3	US-08-469-318-75	Sequence 75, Appl
23	50.6	28.1	921	3	US-08-469-318-78	Sequence 78, Appl
24	50.6	28.1	921	3	US-08-469-318-84	Sequence 84, Appl
25	50.6	28.1	921	3	US-08-468-609A-72	Sequence 72, Appl
26	50.6	28.1	921	3	US-08-468-609A-75	Sequence 75, Appl
27	50.6	28.1	921	3	US-08-468-609A-78	Sequence 78, Appl
28	50.6	28.1	921	3	US-08-468-609A-84	Sequence 84, Appl
29	50.6	28.1	921	4	US-08-446-872A-72	Sequence 72, Appl
30	50.6	28.1	921	4	US-08-446-872A-75	Sequence 75, Appl
31	50.6	28.1	921	4	US-08-446-872A-78	Sequence 78, Appl
32	50.6	28.1	921	4	US-08-446-872A-84	Sequence 84, Appl
33	50.6	28.1	921	4	US-08-762-227A-72	Sequence 72, Appl
34	50.6	28.1	921	4	US-08-762-227A-75	Sequence 75, Appl
35	50.6	28.1	921	4	US-08-762-227A-78	Sequence 78, Appl
36	50.6	28.1	921	4	US-08-762-227A-84	Sequence 84, Appl
37	50.6	28.1	921	5	PCT-US95-01185-72	Sequence 72, Appl
38	50.6	28.1	921	5	PCT-US95-01185-75	Sequence 75, Appl
39	50.6	28.1	921	5	PCT-US95-01185-78	Sequence 78, Appl
40	50.6	28.1	921	5	PCT-US95-01185-84	Sequence 84, Appl
41	50.6	28.1	966	3	US-08-469-318-68	Sequence 68, Appl
42	50.6	28.1	966	3	US-08-469-318-70	Sequence 70, Appl
43	50.6	28.1	966	3	US-08-469-318-71	Sequence 71, Appl
44	50.6	28.1	966	3	US-08-469-318-73	Sequence 73, Appl
45	50.6	28.1	966	3	US-08-469-318-77	Sequence 77, Appl

ALIGNMENTS

```
RESULT 1
US-08-104-445-2
; Sequence 2, Application 08/104445
; Patent No. 5306633
; GENERAL INFORMATION:
;    APPLICANT: GOTTSCHALK, MICHAEL
;    APPLICANT: SCHUSTER, ERWIN
;    APPLICANT: SPROESSLER, BRUNO
;    TITLE OF INVENTION: BACTERIAL XYLANASE, METHOD FOR ITS
;    TITLE OF INVENTION: PRODUCTION, BACTERIA PRODUCING A XYLANASE, DNA FRAGMENT
;    TITLE OF INVENTION: ENCODING A XYLANASE, PLASMID CONTAINING THE DNA FRAGMENT,
;    TITLE OF INVENTION: BAKING AGENTS CONTAINING A XYLANASE, AND METHOD FO
;    NUMBER OF SEQUENCES: 3
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;       ADDRESS:  P.C.
;       STREET: 1755 S. Jefferson Davis Highway, Suite 400
;       CITY: Arlington
;       STATE: Virginia
;       COUNTRY: U.S.A.
;       ZIP: 22202
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: 08/104,445
;       FILING DATE:
;       CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Oblon, No. 5306633man F.
;       REGISTRATION NUMBER:~ 24,618
;       REFERENCE/DOCKET NUMBER: 583-179-0
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (703) 413-3000
;       TELEFAX: (703) 413-2220
;       TELEX: 248855 OPAT UR
;    INFORMATION FOR SEQ ID NO: 2:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 642 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: unknown
;       TOPOLOGY: unknown
;    MOLECULE TYPE: DNA (genomic)
;    FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 1..639
US-08-104-445-2
```

Db 566 ATGTTTCGGCAACCGCCTCTGCAGCTGGCACAGATTACTGGCAAAATTGGACTGGACGGG 625

Qy 121 AGGACTCCGTTAGGTCCAGCCA 142
Db 626 GGGCGGACGCAGTAAACGCAGTCA 647

RESULT 3
US-09-570-856B-2
; Sequence 2, Application US/09570856B
; Patent No. 6682923
; GENERAL INFORMATION:
; APPLICANT: Bentzien, Joerg M
; APPLICANT: Dahiyat, Bassil I
; TITLE OF INVENTION: NOVEL THERMOSTABLE ALKALIPHILIC XYLANASE
; FILE REFERENCE: A-67478-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/570,856B
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/133,714
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1349
; TYPE: DNA
; ORGANISM: Bacillus circulans
US-09-570-856B-2

Query Match   43.9%;   Score 79; DB 4; Length 1349;
Best Local Similarity   79.0%;   Pred. No. 1.6e-16;
Matches   94; Conservative   0; Mismatches   25; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
Db 392 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTAATGAGTATTAGC 451

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCCACCATCACCATATCGAGGG 119
Db 452 TTGTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGG 510

RESULT 4
US-08-575-964-2
; Sequence 2, Application US/08575964
; Patent No. 5736384
; GENERAL INFORMATION:
; APPLICANT: Fukunaga, No. 5736384yuki
; APPLICANT: Iwasaki, Yuji
; APPLICANT: Kono, Satoko
; APPLICANT: Kita, Yukio
; APPLICANT: Izumi, Yoshiya
; TITLE OF INVENTION: THERMOSTABLE XYLANASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson P.C.
;   STREET: 4225 Executive Square, Suite 1400
;   CITY: La Jolla
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 92037
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/575,964
;   FILING DATE: 20-DEC-1995
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Haile, Lisa A.
;   REGISTRATION NUMBER: 38,347

Query Match   49.6%;   Score 89.2; DB 1; Length 642;
Best Local Similarity   76.8%;   Pred. No. 4.8e-20;
Matches 109; Conservative   0; Mismatches   33; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
Db   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCCACCATCACCATATCGAGGGA 120
Db   61 ATGTTTCGGCAACCGCCTCTGCAGCTGGCACAGATTACTGGCAAAATTGGACTGACGGG 120

Qy 121 AGGACTCCGTTAGGTCCAGCCA 142
Db 121 GGGCGGACGCAGTAAACGCAGTCA 142

RESULT 2
US-08-104-445-1
; Sequence 1, Application 08/104445
; Patent No. 5306633
; GENERAL INFORMATION:
; APPLICANT: GOTTSCHALK, MICHAEL
; APPLICANT: SCHUSTER, ERWIN
; APPLICANT: SPROESSLER, BRUNO
; TITLE OF INVENTION: BACTERIAL XYLANASE, METHOD FOR ITS
; TITLE OF INVENTION: PRODUCTION, BACTERIA PRODUCING A XYLANASE, DNA FRAGMENT
; TITLE OF INVENTION: ENCODING A XYLANASE, PLASMID CONTAINING THE DNA FRAGMENT,
; TITLE OF INVENTION: BAKING AGENTS CONTAINING A XYLANASE, AND METHOD FO
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. Jefferson Davis Highway, Suite 400
;   CITY: Arlington
;   STATE: Virginia
;   COUNTRY: U.S.A.
;   ZIP: 22202
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: 08/104,445
;   FILING DATE:
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Oblon, No. 5306633man F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 583-179-0
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703) 413-3000
;   TELEFAX: (703) 413-2220
;   TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1413 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-104-445-1

Query Match   49.6%;   Score 89.2; DB 1; Length 1413;
Best Local Similarity   76.8%;   Pred. No. 6.7e-20;
Matches 109; Conservative   0; Mismatches   33; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
Db 506 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGGATTAACGGCAGCTTTCATGAGTATCAGC 565

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCCACCATCACCATATCGAGGGA 120

RESULT 5
US-08-963-500-2
; Sequence 2, Application US/08963500
; Patent No. 5916795
GENERAL INFORMATION:
APPLICANT: Fukunaga, No. 5916795uyuki
APPLICANT: Iwasaki, Yuji
APPLICANT: Kono, Satoko
APPLICANT: Kita, Yukio
APPLICANT: Izumi, Yoshiya
TITLE OF INVENTION: THERMOSTABLE XYLANASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Fish & Richardson P.C.
  STREET: 4225 Executive Square, Suite 1400
  CITY: La Jolla
  STATE: MA
  COUNTRY: USA
  ZIP: 92037
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/963,500
  FILING DATE:
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/575,964
  FILING DATE: 20-DEC-1995
ATTORNEY/AGENT INFORMATION:
  NAME: Haile, Lisa A.
  REGISTRATION NUMBER: 38,347
  REFERENCE/DOCKET NUMBER: 07898/002001
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 619/678-5070
  TELEFAX: 619/678-5099
  TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 1207 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  ORIGINAL SOURCE:
  ORGANISM: Bacillus sp.
  STRAIN: 2113
  FEATURE:
  NAME/KEY: P CDS
  LOCATION: 379...1029
  IDENTIFICATION METHOD: by experiment
  OTHER INFORMATION:
US-08-575-964-2

Query Match   28.2%; Score 50.8; DB 1; Length 1207;
Best Local Similarity 74.4%; Pred. No. 3.2e-07;
Matches 64; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   1 ATGTTTAAGTTAAAAGAAATTCTTAGTGTGGGATTAACGGCAGCTTCATGAGTATCAGC 60
        ||||| |||||||||||||| |||||| ||||| ||||||||||||||
Db 397 ATGATTAAGTCTAAAAAGAAGAAATTTTTGACGGTATGTATTGCAGCATTAATGAGTTTTAGC 456

Qy 61 ATGTTTTCTGCAACCGCCTCTGCGAGC 86
        ||||| |||||| |||||
Db 457 TTGTTTGCAGCAACCTCAAATGCAGC 482

TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 1207 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  ORIGINAL SOURCE:
  ORGANISM: Bacillus sp.
  STRAIN: 2113
  FEATURE:
  NAME/KEY: P CDS
  LOCATION: 379...1029
  IDENTIFICATION METHOD: by experiment
  OTHER INFORMATION:
US-08-963-500-2

Query Match   28.2%; Score 50.8; DB 2; Length 1207;
Best Local Similarity 74.4%; Pred. No. 3.2e-07;
Matches 64; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   1 ATGTTTAAGTTAAAAGAAATTCTTAGTGTGGGATTAACGGCAGCTTCATGAGTATCAGC 60
        ||||| |||||||||||||| |||||| ||||| ||||||||||||||
Db 397 ATGATTAAGTCTAAAAAGAAGAAATTTTTGACGGTATGTATTGCAGCATTAATGAGTTTTAGC 456

Qy 61 ATGTTTTCTGCAACCGCCTCTGCGAGC 86
        ||||| |||||| |||||
Db 457 TTGTTTGCAGCAACCTCAAATGCAGC 482

RESULT 6
US-08-469-318-177
; Sequence 177, Application US/08469318
; Patent No. 6022535
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 196
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/469,318
  FILING DATE:
  CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/446,872
  FILING DATE:
INFORMATION FOR SEQ ID NO: 177:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 546 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
US-08-469-318-177

Query Match   28.1%; Score 50.6; DB 3; Length 546;
Best Local Similarity 93.0%; Pred. No. 2.7e-07;
Matches 53; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 124 ACTCCGTTAGGTCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
        |||||||||| |||| |||||||||||||||||| |||||| |||||||||||
Db   7 ACACCATTAGGCCCTGCCAGCTCCCTGCCCCAGAGCTCCCTGCTCTCAAGTGCTTAGAG 63

RESULT 7
US-08-469-318-178
; Sequence 178, Application US/08469318

Patent No. 6022535
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 196
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,318
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,872
FILING DATE:
INFORMATION FOR SEQ ID NO: 178:
SEQUENCE CHARACTERISTICS:
LENGTH: 546 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-469-318-178

Query Match   28.1%;   Score 50.6; DB 3;   Length 546;
Best Local Similarity 93.0%; Pred. No. 2.7e-07; Indels 0; Gaps 0;
Matches 53; Conservative 0; Mismatches 4;

Qy 124 ACTCCGTTAGGTCCAGCCAGCCAGAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
        || || |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   7 ACACCATTAGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 8
US-08-468-609A-177
; Sequence 177, Application US/08468609A
; Patent No. 6030812
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Maire H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olins, Peter O.
; APPLICANT: Paik, Kunnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Inteleukin-3 (IL-
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,609A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192,325
; FILING DATE: 14-FEB-1994

; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2790/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6986
; TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 177:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-468-609A-177

Query Match   28.1%;   Score 50.6; DB 3;   Length 546;
Best Local Similarity 93.0%; Pred. No. 2.7e-07; Indels 0; Gaps 0;
Matches 53; Conservative 0; Mismatches 4;

Qy 124 ACTCCGTTAGGTCCAGCCAGCCAGAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
        || || |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   7 ACACCATTAGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 9
US-08-468-609A-178
; Sequence 178, Application US/08468609A
; Patent No. 6030812
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Maire H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olins, Peter O.
; APPLICANT: Paik, Kunnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Inteleukin-3 (IL-
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,609A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192,325
; FILING DATE: 14-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2790/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6986
; TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 base pairs
; TYPE: nucleic acid

```
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
US-08-468-609A-178

Query Match          28.1%;   Score 50.6;  DB 3;  Length 546;
Best Local Similarity 93.0%;   Pred. No. 2.7e-07;
Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy    124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
          |||||||||||  | |||||||||||||||||||||||||||||||||||||||||
Db      7 ACACCATTAGGACCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 10
US-08-446-872A-177
; Sequence 177, Application US/08446872A
; Patent No. 6361977
; GENERAL INFORMATION:
;      APPLICANT: Abrams, Mark A.
;      APPLICANT: Bauer, S. C.
;      APPLICANT: Braford-Goldberg, Sarah R.
;      APPLICANT: Caparon, Maire H.
;      APPLICANT: Easton, Alan M.
;      APPLICANT: Klein, Barbara K.
;      APPLICANT: McKearn, John P.
;      APPLICANT: Olins, Peter O.
;      APPLICANT: Paik, Kumnan
;      APPLICANT: Thomas, John W.
;      TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis
;      TITLE OF INVENTION: Fusion Protein
;      NUMBER OF SEQUENCES: 197
;      CORRESPONDENCE ADDRESS:
;         ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
;         ADDRESSEE: Corporate Patent Dept.
;         STREET: P. O. Box 5110
;         CITY: Chicago
;         STATE: Illinois
;         COUNTRY: USA
;         ZIP: 60680
;      COMPUTER READABLE FORM:
;         MEDIUM TYPE: Floppy disk
;         COMPUTER: IBM PC compatible
;         OPERATING SYSTEM: PC-DOS/MS-DOS
;         SOFTWARE: PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;         APPLICATION NUMBER: US/08/446,872A
;         FILING DATE: 06-JUN-1995
;         CLASSIFICATION: 424
;      PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: US 08/192,325
;         FILING DATE: 14-FEB-1994
;      ATTORNEY/AGENT INFORMATION:
;         NAME: Bennett, Dennis A.
;         REGISTRATION NUMBER: 34,547
;         REFERENCE/DOCKET NUMBER: C-2790/1
;      TELECOMMUNICATION INFORMATION:
;         TELEPHONE: (314)737-6986
;         TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 177:
;      SEQUENCE CHARACTERISTICS:
;         LENGTH: 546 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: double
;         TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
US-08-446-872A-177

Query Match          28.1%;   Score 50.6;  DB 4;  Length 546;
Best Local Similarity 93.0%;   Pred. No. 2.7e-07;
Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy    124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
          |||||||||||  | |||||||||||||||||||||||||||||||||||||||||
Db      7 ACACCATTAGGACCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 11
US-08-446-872A-178
; Sequence 178, Application US/08446872A
; Patent No. 6361977
; GENERAL INFORMATION:
;      APPLICANT: Abrams, Mark A.
;      APPLICANT: Bauer, S. C.
;      APPLICANT: Braford-Goldberg, Sarah R.
;      APPLICANT: Caparon, Maire H.
;      APPLICANT: Easton, Alan M.
;      APPLICANT: Klein, Barbara K.
;      APPLICANT: McKearn, John P.
;      APPLICANT: Olins, Peter O.
;      APPLICANT: Paik, Kumnan
;      APPLICANT: Thomas, John W.
;      TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis
;      TITLE OF INVENTION: Fusion Protein
;      NUMBER OF SEQUENCES: 197
;      CORRESPONDENCE ADDRESS:
;         ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
;         ADDRESSEE: Corporate Patent Dept.
;         STREET: P. O. Box 5110
;         CITY: Chicago
;         STATE: Illinois
;         COUNTRY: USA
;         ZIP: 60680
;      COMPUTER READABLE FORM:
;         MEDIUM TYPE: Floppy disk
;         COMPUTER: IBM PC compatible
;         OPERATING SYSTEM: PC-DOS/MS-DOS
;         SOFTWARE: PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;         APPLICATION NUMBER: US/08/446,872A
;         FILING DATE: 06-JUN-1995
;         CLASSIFICATION: 424
;      PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: US 08/192,325
;         FILING DATE: 14-FEB-1994
;      ATTORNEY/AGENT INFORMATION:
;         NAME: Bennett, Dennis A.
;         REGISTRATION NUMBER: 34,547
;         REFERENCE/DOCKET NUMBER: C-2790/1
;      TELECOMMUNICATION INFORMATION:
;         TELEPHONE: (314)737-6986
;         TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 178:
;      SEQUENCE CHARACTERISTICS:
;         LENGTH: 546 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: double
;         TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
US-08-446-872A-178

Query Match          28.1%;   Score 50.6;  DB 4;  Length 546;
Best Local Similarity 93.0%;   Pred. No. 2.7e-07;
Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy    124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
          |||||||||||  | |||||||||||||||||||||||||||||||||||||||||
Db      7 ACACCATTAGGACCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 12
US-08-762-227A-177
; Sequence 177, Application US/08762227A
; Patent No. 6436387
; GENERAL INFORMATION:
;      APPLICANT: Abrams, Mark A.
```

                    Paik, Kumnan
                    Thomas, John W.
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis
                    Fusion Protein
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
    ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
               Corporate Patent Dept.
    STREET: P. O. Box 5110
    CITY: Chicago
    STATE: Illinois
    COUNTRY: USA
    ZIP: 60680
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/762,227A
    FILING DATE: 09-Dec-1996
    CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/192,325
    FILING DATE: 14-FEB-1994
    APPLICATION NUMBER: US 08/446,872
    FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
    NAME: Bennett, Dennis A.
    REGISTRATION NUMBER: 34,547
    REFERENCE/DOCKET NUMBER: C-2790/5
TELECOMMUNICATION INFORMATION:
    TELEPHONE: (708)470-6501
    TELEFAX: (708)470-6881
INFORMATION FOR SEQ ID NO: 178:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 546 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-08-762-227A-178

Query Match   28.1%; Score 50.6; DB 4; Length 546;
Best Local Similarity 93.0%; Pred. No. 2.7e-07;
Matches 53; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 ACACCATTAGGACCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 14
PCT-US95-01185-177
; Sequence 177, Application PC/TUS9501185
; GENERAL INFORMATION:
    APPLICANT:
    TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
    TITLE OF INVENTION: Protein
    NUMBER OF SEQUENCES: 196
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: PCT/US95/01185
    FILING DATE: 02-FEB-1995
    CLASSIFICATION:
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/192325
    FILING DATE: 14-FEB-1994

```
;    INFORMATION FOR SEQ ID NO: 177:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  546 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  double
;        TOPOLOGY:  linear
;      MOLECULE TYPE:  DNA (genomic)
PCT-US95-01185-177

Query Match             28.1%;  Score 50.6;  DB 5;  Length 546;
Best Local Similarity   93.0%;  Pred. No. 2.7e-07;
Matches   53;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        124 ACTCCGTTAGGTCCAGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
              || ||| |||||| ||||||||||||||||||||||||||||||||||||||||||
Db          7 ACACCATTAGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63


RESULT 15
PCT-US95-01185-178
; Sequence 178, Application PC/TUS9501185
; GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Multivariant IL-3 Hematopoiesis Fusion
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  196
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/01185
;      FILING DATE:  02-FEB-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/192325
;      FILING DATE:  14-FEB-1994
;    INFORMATION FOR SEQ ID NO: 178:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  546 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  double
;        TOPOLOGY:  linear
;      MOLECULE TYPE:  DNA (genomic)
PCT-US95-01185-178

Query Match             28.1%;  Score 50.6;  DB 5;  Length 546;
Best Local Similarity   93.0%;  Pred. No. 2.7e-07;
Matches   53;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        124 ACTCCGTTAGGTCCAGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
              || ||| |||||| ||||||||||||||||||||||||||||||||||||||||||
Db          7 ACACCATTAGGACCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63


Search completed: October  9, 2004, 04:24:24
Job time : 28.6529 secs
```

This Page Blank (uspto)

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2004, 21:19:03 ; Search time 136.728 Seconds
                                    (without alignments)
                            5592.665 Million cell updates/sec

Title: US-10-009-792C-26
Perfect score: 180
Sequence: 1 atgttaagttaaaagaa.........tcctgctcaagtgcttagag 180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
        1: geneseqn1980s:*
        2: geneseqn1990s:*
        3: geneseqn2000s:*
        4: geneseqn2001as:*
        5: geneseqn2001bs:*
        6: geneseqn2002s:*
        7: geneseqn2003as:*
        8: geneseqn2003bs:*
        9: geneseqn2003cs:*
       10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	180	100.0	180	4	AAD19775	Aad19775 Human hG-
2	89.2	49.6	642	2	AAQ56052	Aaq56052 Sequence
3	89.2	49.6	1413	2	AAQ66480	Aaq66480 Sequence
4	87.6	48.7	642	5	AAA47153	Aaa47153 DNA encod
5	86	47.8	135	4	AAD19774	Aad19774 Human hG-
6	79.6	44.2	642	3	AAA47157	Aaa47157 DNA encod
7	79.6	44.2	642	3	AAA47155	Aaa47155 DNA encod
8	79.6	44.2	642	3	AAA47156	Aaa47156 DNA encod
9	79	43.9	642	3	AAA47154	Aaa47154 DNA encod
10	79	43.9	645	5	AAS13814	Aas13814 DNA encod
11	79	43.9	645	3	ADC27538	Adc27538 Bacillus
12	79	43.9	1349	6	AAC88889	Aac88889 Bacillus
13	79	43.9	1349	6	AAL48894	Aal48894 B circula
14	52	28.9	644	5	AAI71848	Aai71848 Recombina
15	50.6	28.1	546	2	AAQ97205	Aaq97205 pMON13010
16	50.6	28.1	546	2	AAQ97206	Aaq97206 pMON13499
17	50.6	28.1	546	3	AAA03772	Aaa03772 Human G-C
18	50.6	28.1	546	3	AAA03773	Aaa03773 Human G-C
19	50.6	28.1	546	6	ABX00088	Abx00088 Human int
20	50.6	28.1	546	6	ABX00087	Abx00087 Human int
21	50.6	28.1	630	1	AAN70029	Aan70029 Mature na
22	50.6	28.1	906	2	AAQ97167	Aaq97167 pMON13023
23	50.6	28.1	906	3	AAA03721	Aaa03721 Human int
24	50.6	28.1	906	6	ABX00010	Abx00010 Human int
25	50.6	28.1	921	2	AAQ97198	Aaq97198 pMON13066
26	50.6	28.1	921	2	AAQ97186	Aaq97186 pMON13026
27	50.6	28.1	921	2	AAQ97189	Aaq97189 pMON13043
28	50.6	28.1	921	3	AAQ97192	Aaq97192 pMON13151
29	50.6	28.1	921	3	AAA03746	Aaa03746 Human int
30	50.6	28.1	921	3	AAA03752	Aaa03752 Human int
31	50.6	28.1	921	3	AAA03740	Aaa03740 Human int
32	50.6	28.1	921	3	AAA03743	Aaa03743 Human int
33	50.6	28.1	921	6	ABX00029	Abx00029 Human int
34	50.6	28.1	921	6	ABX00035	Abx00035 Human int
35	50.6	28.1	921	6	ABX00032	Abx00032 Human int
36	50.6	28.1	921	6	ABX00041	Abx00041 Human int
37	50.6	28.1	966	2	AAQ97185	Aaq97185 pMON13060
38	50.6	28.1	966	2	AAQ97191	Aaq97191 pMON13045
39	50.6	28.1	966	2	AAQ97184	Aaq97184 pMON13058
40	50.6	28.1	966	2	AAQ97187	Aaq97187 pMON13063
41	50.6	28.1	966	2	AAQ97193	Aaq97193 pMON13152
42	50.6	28.1	966	2	AAQ97182	Aaq97182 pMON13034
43	50.6	28.1	966	3	AAA03741	Aaa03741 Human int
44	50.6	28.1	966	3	AAA03745	Aaa03745 Human int
45	50.6	28.1	966	3	AAA03736	Aaa03736 Human int

ALIGNMENTS

RESULT 1
AAD19775
ID   AAD19775 standard; DNA; 180 BP.
XX
AC   AAD19775;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human hG-CSF DNA 5' end inserted into plasmid pTHSCSFmII.
XX
KW   Human; granulocyte colony stimulating factor; hG-CSF; protease;
KW   Factor Xa; kanamycin resistance; endoxylanase signal peptide; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..180
FT   /*tag= a
FT   /product= "Human hG-CSF protein fragment"
FT   /note= "CDS does not include stop codon"
FT   /partial
FT   sig_peptide   1..84
FT   /*tag= b
FT   /note= "Endoxylanase signal sequence"
FT   misc_feature   85..123
FT   /*tag= c
FT   /note= "Nucleotides encoding an inserted oligopeptide"
FT   mat_peptide   124..180
FT   /*tag= d
FT   /product= "Human mature hG-CSF protein fragment"
XX
PN   WO200173081-A1.
XX
PD   04-OCT-2001.
XX
PF   31-MAR-2001; 2001WO-KR000549.
XX
PR   31-MAR-2000; 2000XR-00017052.
XX
PA   (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI   Lee S, Jeong K;
XX
DR   WPI; 2001-616523/71.
DR   P-PSDB; AAB12157.
XX

PT   Recombinant plasmid vector comprising an endoxylanase signal sequence,
PT   human granulocyte colony stimulating factor gene and other components,
PT   when transformed into microorganism useful for preparing the colony
PT   stimulating factor.
XX
PS   .Claim 3; Fig 12; 50pp; English.
XX
CC   The invention relates to an Escherichia coli producing and secreting
CC   human granulocyte colony stimulating factor (hG-CSF), more specifically,
CC   to a recombinant plasmid constructed to express secretory hG-CSF in E.
CC   coli, an E. coli transformed with that plasmid to secrete hG-CSF, and a
CC   process for preparing hG-CSF using the transformed hG-CSF. The
CC   recombinant plasmid vector comprises of a kanamycin resistance gene, a
CC   promoter, an endoxylanase signal sequence, a nucleotide sequence coding
CC   for an oligopeptide consisting of 13 amino acids including 6 consecutive
CC   histidine residues and a hG-CSF. E.coli transformed with recombinant
CC   plasmid vector is useful for preparing hG-CSF. The method comprises
CC   culturing the microorganism to obtain a hG-CSF fusion protein and
CC   treating the fusion protein with a protease preferably Factor Xa, to
CC   obtain a hG-CSF, where the fusion protein is obtained from the culture by
CC   employing Ni-column. The present sequence is a DNA encoding human hG-CSF
CC   N-terminal region inserted into plasmid pTHsCSFmII
XX
SQ   Sequence 180 BP; 44 A; 50 C; 40 G; 46 T; 0 U; 0 Other;

Query Match   100.0%; Score 180; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-50;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGGCCCGCCACCATCACCATCACCATATCGAGGGA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGGCCCGCCACCATCACCATCACCATATCGAGGGA 120

Qy 121 AGGACTCCGTTAGGTCCAGCCAGCTCCCTGCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGGACTCCGTTAGGTCCAGCCAGCTCCCTGCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180

RESULT 2
AAQ56052
ID   AAQ56052 standard; DNA; 642 BP.
XX
AC   AAQ56052;
XX
DT   25-MAR-2003 (revised)
DT   10-AUG-1994 (first entry)
XX
DE   Sequence encoding xylanase.
XX
KW   Xylanase; Bacillus subtilis; baking; bread; yeast; food; foodstuffs;
KW   bread; pastry; ss.
XX
OS   Bacillus subtilis.
XX
PN   DE4226528-A1.
XX
PD   17-FEB-1994.
XX
PF   11-AUG-1992; 92DE-04226528.
XX
PR   11-AUG-1992; 92DE-04226528.
XX
PA   (ROHG) ROEHM GMBH.
XX
PI   Gottschalk M, Sproessler B, Schuster E;
XX
DR   WPI; 1994-058089/08.
DR   P-PSDB; AAR47200.
XX

PT   New xylanase obtd. from Bacillus subtilis - useful in baking agents for
PT   increased vol. of baking prods.
XX
PS   Claim 4; Page 10; 11pp; German.
XX
CC   The modified xylanase is used as a baking agent, preferably for yeast
CC   pastry and white bread. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 642 BP; 191 A; 122 C; 166 G; 163 T; 0 U; 0 Other;

Query Match   49.6%; Score 89.2; DB 2; Length 642;
Best Local Similarity 76.8%; Pred. No. 1.2e-19;
Matches 109; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy   1 ATGTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCCACCATCACCATATCGAGGGA 120
         |||||||||||||||||||||||||||||||||| | | | |||
Db   61 ATGTTTCCGCAACCGCCTCTGCAGCTGGCCCACAGATTACTGCGCAAATTGGACGACGGG 120

Qy 121 AGGACTCCGTTAGGTCCAGCCA 142
         |||||||||
Db 121 GGCGGGACAGTAAACGCAGTCA 142

RESULT 3
AAQ66480
ID   AAQ66480 standard; DNA; 1413 BP.
XX
AC   AAQ66480;
XX
DT   25-MAR-2003 (revised)
DT   10-AUG-1994 (first entry)
XX
DE   Sequence comprising xylanase coding sequence.
XX
KW   Xylanase; Bacillus subtilis; baking; bread; yeast; food; foodstuffs;
KW   bread; pastry; ss.
XX
OS   Bacillus subtilis.
XX
FH   Key   Location/Qualifiers
FT   CDS   506..1209
FT   /*tag= a
FT   /product= "Modified xylanase."
XX
PN   DE4226528-A1.
XX
PD   17-FEB-1994.
XX
PF   11-AUG-1992; 92DE-04226528.
PR   11-AUG-1992; 92DE-04226528.
XX
PA   (ROHG) ROEHM GMBH.
XX
PI   Gottschalk M, Sproessler B, Schuster E;
XX
DR   WPI; 1994-058089/08.
XX
PT   New xylanase obtd. from Bacillus subtilis - useful in baking agents for
PT   increased vol. of baking prods.
XX
PS   Example 4; Page 9; 11pp; German.
XX
CC   The modified xylanase is used as a baking agent, preferably for yeast
CC   pastry and white bread. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 1413 BP; 408 A; 280 C; 317 G; 408 T; 0 U; 0 Other;

Query Match   49.6%; Score 89.2; DB 2; Length 1413;

```
Query Match          48.7%;  Score 87.6;  DB 3;  Length 642;
Best Local Similarity 76.1%;  Pred. No. 4.2e-19;
Matches 108;  Conservative 0;  Mismatches 34;  Indels 0;  Gaps 0;

Qy   1 ATGTTTAAGTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC  60
        |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1 ATGTTTAAGTTAAAAGAAATTCTTAGTGGATTAACGGCAGCTTTCATGAGTATCAGC   60

Qy  61 ATGTTTTCGCAACCGCCCTCTCGAGCTGGCCGCACCATCACCATCACCATATCGAGGGA 120
        |||||||| ||||||||  ||||||||||||| |  |     || | |  ||
Db  61 ATGTTTTCGGCAACCGCCGCCTCTCGAGCTGGCCACAGATTACTGGCAAAATTGGACTGACGGG 120

Qy 121 AGGACTCCGTTAGGTCCAGCCA 142
        ||| ||   ||    || ||
Db 121 GGCGGGACCAGTAAACGCAGTCA 142
```

RESULT 5
AAD19774
```
ID   AAD19774 standard; DNA; 135 BP.
XX
AC   AAD19774;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human hG-CSF DNA 5' end inserted into plasmid pTrcSCSFmII.
XX
KW   Human; granulocyte colony stimulating factor; hG-CSF; protease;
KW   Factor Xa; kanamycin resistance; endoxylanase signal peptide; ds.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..135
FT                  /*tag= a
FT                  /product= "Human hG-CSF protein fragment"
FT                  /note= "CDS does not include stop codon"
FT                  /partial
FT   sig_peptide    1..84
FT                  /*tag= b
FT   mat_peptide    85..135
FT                  /*tag= c
FT                  /product= "Human mature hG-CSF protein fragment"
XX
PN   WO200173081-A1.
XX
PD   04-OCT-2001.
XX
PF   31-MAR-2001; 2001WO-KR000549.
XX
PR   31-MAR-2000; 2000KR-00017052.
XX
PA   (KOAD ) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI   Lee S, Jeong K;
XX
DR   WPI; 2001-616523/71.
DR   P-PSDB; AAE12156.
XX
PT   Recombinant plasmid vector comprising an endoxylanase signal sequence,
PT   human granulocyte colony stimulating factor gene and other components,
PT   when transformed into microorganism useful for preparing the colony
PT   stimulating factor.
XX
PS   Example 4; Fig 9; 50pp; English.
XX
CC   The invention relates to an Escherichia coli producing and secreting
CC   human granulocyte colony stimulating factor (hG-CSF), more specifically,
CC   to a recombinant plasmid constructed to express secretory hG-CSF in E.
CC   coli, an E. coli transformed with that plasmid to secrete hG-CSF, and a
CC   process for preparing hG-CSF using the transformed hG-CSF. The
CC   recombinant plasmid vector comprises of a kanamycin resistance gene, a
CC   promoter, an endoxylanase signal sequence, a nucleotide sequence coding
```

```
Best Local Similarity 76.8%;  Score 87.6;  Pred. No. 1.7e-19;
Matches 109;  Conservative 0;  Mismatches 33;  Indels 0;  Gaps 0;

Qy   1 ATGTTTAAGTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 506 ATGTTTAAGTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 565

Qy  61 ATGTTTTCGCAACCGCCCTCTCGAGCTGGCCGCACCATCACCATCACCATATCGAGGGA 120
        |||||||| ||||||||  ||||||||||||| |  |     || | |  ||
Db 566 ATGTTTTCGGCAACCGCCGCCTCTCGAGCTGGCCACAGATTACTGGCAAAATTGGACTGACGGG 625

Qy 121 AGGACTCCGTTAGGTCCAGCCA 142
        ||| ||   ||    || ||
Db 626 GGCGGGACAGTAAACGCAGTCA 647
```

RESULT 4
AAA47153
```
ID   AAA47153 standard; DNA; 642 BP.
XX
AC   AAA47153;
XX
DT   03-OCT-2000  (first entry)
XX
DE   DNA encoding a bacterial endo-beta-1,4-xylanase protein.
XX
KW   Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
KW   dough; dough preparation; ss.
XX
OS   Bacillus subtilis.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..642
FT                  /*tag= a
FT                  /product= "xylanase"
XX
PN   WO200039289-A2.
XX
PD   06-JUL-2000.
XX
PF   17-DEC-1999; 99WO-IB002071.
XX
PR   23-DEC-1998; 98GB-00028599.
PR   06-APR-1999; 99GB-00007805.
PR   15-APR-1999; 99GB-0008645.
XX
PA   (DANI-) DANISCO AS.
XX
PI   Sibbesen O, Sorensen JF;
XX
DR   WPI; 2000-465744/40.
DR   P-PSDB; AAY93751.
XX
PT   Mutant xylanase protein identified using xylanase inhibitor useful for
PT   preparing non-sticky dough for bakery products.
XX
PS   Disclosure; Page 107; 112pp; English.
XX
CC   The present sequence encodes an endo-beta-1,4-xylanase. The specification
CC   also describes an endo-beta-1,4-xylanase inhibitor, which is obtained
CC   from wheat flour. The specification also describes a mutant xylanase
CC   protein. The xylanase is useful for preparing a foodstuff, preferably a
CC   bakery product or a substance (e.g. a dough) for making the bakery
CC   product. Wild type xylanase or mutant xylanase is useful for preparing a
CC   dough that is less sticky than a dough comprising a fungal xylanase. The
CC   xylanase inhibitor is useful for screening high degree resistance
CC   xylanases for dough preparation. The xylanase is also useful for
CC   preparing a non-sticky dough. A combination of xylanase and the inhibitor
CC   is useful for calibrating and/or determining the quantity of inhibitor in
CC   a wheat flour sample
XX
SQ   Sequence 642 BP; 192 A; 122 C; 164 G; 164 T; 0 U; 0 Other;
```

CC   for an oligopeptide consisting of 13 amino acids including 6 consecutive
CC   histidine residues and a hG-CSF. E.coli transformed with recombinant
CC   plasmid vector is useful for preparing hG-CSF. The method comprises
CC   culturing the microorganism to obtain a hG-CSF fusion protein and
CC   treating the fusion protein with a protease preferably Factor Xa, to
CC   obtain a hG-CSF, where the fusion protein is obtained from the culture by
CC   employing Ni-column. The present sequence is a DNA encoding human hG-CSF
CC   N-terminal region inserted into plasmid pTrcCSFmII
XX
SQ   Sequence 135 BP; 32 A; 36 C; 28 G; 39 T; 0 U; 0 Other;

Query Match   47.8%; Score 86; DB 4; Length 135;
Best Local Similarity   77.6%; Pred. No. 7.8e-19;
Matches 135; Conservative   0; Mismatches   0; Indels   39; Gaps   1;

Qy   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCGCCGCACCATCACCATCACCATATCGAGGGA 120
              |||||||||||||||||||||
Db   61 ATGTTTCTGCAACCGCCTCTGCA------------------------------------- 84

Qy   121 AGGACTCCGTTAGTCCAGCCAGCTTCCCTGCCCCAGAGCTTCCTGCTCAAGTGC 174
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 ---ACTCCGTTAGTCCAGCCAGCTTCCCTGCCCCAGAGCTTCCTGCTCAAGTGC 135

RESULT 6
AAA47157
ID   AAA47157 standard; DNA; 642 BP.
XX
AC   AAA47157;
XX
DT   03-OCT-2000 (first entry)
XX
DE   DNA encoding a bacterial endo-beta-1,4-xylanase mutant XM3.
XX
KW   Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
KW   dough; dough preparation; ss.
XX
OS   Synthetic.
OS   Bacillus subtilis.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..642
FT   /*tag= a
FT   /product= "xylanase mutant"
XX
PN   WO200039289-A2.
XX
PD   06-JUL-2000.
XX
PF   17-DEC-1999;   99WO-IB002071.
XX
PR   23-DEC-1998;   98GB-00028599.
PR   06-APR-1999;   99GB-00007805.
PR   15-APR-1999;   99GB-00008645.
XX
PA   (DANI-) DANISCO AS.
XX
PI   Sibbesen O, Sorensen JF;
XX
DR   WPI; 2000-465744/40.
DR   P-PSDB; AAY93755.
XX
PT   Mutant xylanase protein identified using xylanase inhibitor useful for
PT   preparing non-sticky dough for bakery products.
XX
PS   Claim 3; Page 111; 112pp; English.
XX
CC   The present sequence encodes a mutant endo-beta-1,4-xylanase. The
CC   specification also describes an endo-beta-1,4-xylanase inhibitor, which

CC   is obtained from wheat flour. The specification also describes a mutant
CC   xylanase protein. The xylanase is useful for preparing a foodstuff,
CC   preferably a bakery product or a substance (e.g. a dough) for making the
CC   bakery product. Wild type xylanase or mutant xylanase is useful for
CC   preparing a dough that is less sticky than a dough comprising a fungal
CC   xylanase. The xylanase inhibitor is useful for screening high degree
CC   resistance xylanases for dough preparation. The xylanase is also useful
CC   for preparing a non-sticky dough. A combination of xylanase and the
CC   inhibitor is useful for calibrating and/or determining the quantity of
CC   inhibitor in a wheat flour sample
XX
SQ   Sequence 642 BP; 192 A; 110 C; 162 G; 178 T; 0 U; 0 Other;

Query Match   44.2%; Score 79.6; DB 3; Length 642;
Best Local Similarity   72.5%; Pred. No. 2.1e-16;
Matches 103; Conservative   0; Mismatches   39; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
              ||||||||||||||||||||||||||| ||||||| |||||||||||||||||||||||
Db   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTTTAATGAGTATTAGC 60

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCGCCGCACCATCACCATCACCATATCGAGGGA 120
               ||| |||||||||||||| |||| ||| ||| | ||| | |||| ||| ||| |
Db   61 TTGTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGCCAAAATTGGACTGATGGG 120

Qy   121 AGGACTCCGTTAGTCCAGCCA 142
                ||| ||||||| |||
Db   121 GGCGGTACCGTAAACGCTGTCA 142

RESULT 7
AAA47155
ID   AAA47155 standard; DNA; 642 BP.
XX
AC   AAA47155;
XX
DT   03-OCT-2000 (first entry)
XX
DE   DNA encoding a bacterial endo-beta-1,4-xylanase mutant XM1.
XX
KW   Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
KW   dough; dough preparation; ss.
XX
OS   Synthetic.
OS   Bacillus subtilis.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..642
FT   /*tag= a
FT   /product= "xylanase mutant"
XX
PN   WO200039289-A2.
XX
PD   06-JUL-2000.
XX
PF   17-DEC-1999;   99WO-IB002071.
XX
PR   23-DEC-1998;   98GB-00028599.
PR   06-APR-1999;   99GB-00007805.
PR   15-APR-1999;   99GB-00008645.
XX
PA   (DANI-) DANISCO AS.
XX
PI   Sibbesen O, Sorensen JF;
XX
DR   WPI; 2000-465744/40.
DR   P-PSDB; AAY93753.
XX
PT   Mutant xylanase protein identified using xylanase inhibitor useful for
PT   preparing non-sticky dough for bakery products.
XX
PS   Claim 3; Page 109; 112pp; English.
XX

This Page Blank (uspto)

XX
SQ   Sequence 546 BP; 95 A; 189 C; 157 G; 105 T; 0 U; 0 Other;

      Query Match   28.1%;   Score 50.6;   DB 2;   Length 546;
      Best Local Similarity   93.0%;   Pred. No. 1.2e-06;
      Matches   53;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;

Qy   124 ACTCCGTTAGTCCAGCAGCTCCCTGCCCAGAGCTTCTGCTCAAGTGCTTAGAG 180
            ||| |||||| || ||||||||||||||||||||||||||||||||||||||
Db   7 ACACCATTAGGACCTGCCAGCTCCCTGCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

Search completed: October 9, 2004, 02:07:18
Job time : 138.728 secs

CC   protein has been modified to exhibit enhanced thermophilicity,
CC   alkalophilicity, or thermostability relative to naturally occurring B.
CC   circulans xylanase, and has at least 5 amino acid substitutions. A
CC   bleaching agent comprising a modified xylanase is useful for bleaching
CC   pulp, in the bioconversion of lignocellulosic materials to fuels, for
CC   clarifying juice and wine, extracting coffee, plant oils and starch,
CC   producing food thickeners, altering texture in bakery products, e.g.
CC   improving the quality of dough, helping bread to rise and processing of
CC   wheat and corn for starch production, use as animal food additives to aid
CC   in the digestibility of feedstuffs and in the washing of super precision
CC   devices and semiconductors. The present sequence is a coding sequence of
CC   a xylanase protein described in the exemplification of the invention
XX
SQ   Sequence 1349 BP; 396 A; 232 C; 312 G; 409 T; 0 U; 0 Other;

        Query Match   43.9%; Score 79; DB 6; Length 1349;
        Best Local Similarity 79.0%; Pred. No. 4.5e-16;
        Matches 94; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTCATGAGTATCAGC 60
        |||||||||||||||||||||||| |||| |||||||| |||||||||| |||||||||
Db 392 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGATTATGGCAGCTTTAATGAGTATTAGC 451

Qy   61 ATGTTTTCTGCAACCGCCGCCTGCGCCCGCACCATCACCATCACCATATGCGAGG 119
        ||||||| ||||||||| | |||| |||| |||| |||| ||||| ||||||
Db 452 TTGTTTTCGGCAACCGCCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGG 510

RESULT 14
AAI71848
ID   AAI71848 standard; cDNA; 644 BP.
XX
AC   AAI71848;
XX
DT   07-JAN-2002 (first entry)
XX
DE   Recombinant human granulocyte colony stimulating factor cDNA.
XX
KW   Human; granulocyte colony stimulating factor; rhG-CSF; recombinant; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   94..621
FT   /*tag= a
FT   /product= "rhG-CSF"
XX
PN   CN167150-A.
XX
PR   CN167150-A.
PD   10-DEC-1997.
XX
PF   05-JUN-1996;   96CN-0106418.
XX
PR   05-JUN-1996;   96CN-0106418.
XX
PA   (JIUY-) JIUYUAN GENE ENG CO LTD HANGZHOU.
PI   Su Y, Kong T, Wang C;
XX
DR   WPI; 2001-590346/67.
DR   P-PSDB; AAM51536.
XX
PT   Production of recombined human granulocyte colony stimulation factor.
XX
PS   Disclosure; Page 8 (disclosure); 15pp; Chinese.
XX
CC   The invention relates to a method for producing recombinant human
CC   granulocyte colony stimulating factor (rhG-CSF). The method includes the
CC   following steps: (a) reverse transcription-polymerase chain reaction of
CC   the human granulocyte colony stimulating factor gene; (b) transforming
CC   Escherichia coli; (c) renaturation of protein by hollow fibre
CC   ultrafiltration dialysis; (d) passing the renatured protein through ion
CC   exchange chromatography, hydrophobic chromatography and molecular sieve

CC   chromatography which are combined together sequentially; and (e)
CC   purifying so to obtain a high yield of high-purity medicinal rhG-CSF
CC   protein. The present sequence encodes rhG-CSF
XX
SQ   Sequence 644 BP; 128 A; 216 C; 172 G; 128 T; 0 U; 0 Other;

        Query Match   28.9%; Score 52; DB 5; Length 644;
        Best Local Similarity 91.7%; Pred. No. 4.2e-07;
        Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 121 AGGACTCCGTTAGGTCCAGCCAGCTCCTCGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
        || || |||||| |||| || || || ||||||||||||||||||||||||
Db   94 ATGACACCATTAGGCCCTGCCAGCTTCCTGCTCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 153

RESULT 15
AAQ97205
ID   AAQ97205 standard; DNA; 546 BP.
XX
AC   AAQ97205;
XX
DT   25-AUG-1999 (first entry)
XX
DE   pMON13010 DNA sequence.
XX
KW   Interleukin; hIL-3; CSF; colony stimulating factor; cytokine; lymphokine;
KW   mutant; mutein; fusion protein; linker; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9521254-A1.
XX
PD   10-AUG-1995.
XX
PF   02-FEB-1995;   95WO-US001185.
XX
PR   04-FEB-1994;   94US-00192325.
XX
PA   (SEAR) SEARLE & CO G D.
XX
PI   Bauer CS,   Abrams MA,   Braford-Goldberg SR,   Caparon MH,   Easton AM;
PI   Klein BK,   Mckearn JP,   Olins PO,   Paik K,   Thomas JW;
XX
DR   WPI; 1995-283774/37.
DR   P-PSDB; AAR79336.
XX
PT   Fusion proteins comprising a human interleukin-3 variant, a linker and
PT   interleukin-3, a variant or a colony stimulating factor - useful to
PT   increase haematopoietic cell prodn. in a mammal.
XX
PS   Example 15; Page 183-184; 447pp; English.
XX
CC   A new fusion protein is disclosed which has the formula R1-L-R2, R2-L-R1,
CC   R1-R2, R2-R1, R1-L-R1 or R1-R1, where R1 is a mutant or variant of human
CC   interleukin-3 (hIL-3), R2 is a second colony stimulating factor (CSF)
CC   including cytokine, lymphokine, interleukin, haematopoietic growth factor
CC   or IL-3 variant, and L is a linker. Generic sequences are described in
CC   AAW03235 - AAW03242, and specifically claimed examples are shown in
CC   AAR79298-R79335 and AAR79342-R79345. The fusion protein is made by
CC   recombinant DNA techniques. Specifically claimed examples of DNA
CC   sequences which encode these proteins are shown in AAQ97167-Q97204 and
CC   AAQ97222-Q97227. The fusion protein is used to increase haematopoietic
CC   cell production. It is also useful as an IL-3 antagonist or as a discrete
CC   antigenic fragment for production of antibodies useful in immunoassays
CC   and immunotherapy. Antagonists are used to block the growth of certain
CC   cancer cells and in treatment of asthma. The fusion protein can also be
CC   used to stimulate bone marrow and blood cell activation and growth in
CC   vitro before infusion; and to treat diseases characterised by decreased
CC   levels of myeloid, erythroid, lymphoid and/or megakaryocyte cells of the
CC   haematopoietic system. The protein has the usual activity of both its
CC   component proteins, but may have increased synergistic activity and
CC   reduced undesired side effects

DR   P-PSDB; ADC27541, ADC27530.
XX
PT   Novel variant xylanase polypeptide having altered sensitivity to a
PT   xylanase inhibitor and altered thermosensitivity as compared with a
PT   parent xylanase enzyme, useful in baking, processing cereals, starch
PT   production.
XX
PS   Example 2; Page 38; 63pp; English.
XX
CC   This invention relates to a novel variant xylanase protein or its
CC   fragment having xylanase activity. the variants have one or more amino
CC   acid modifications so that the protein or fragment has an altered
CC   sensitivity to a xylanase inhibitor and has an altered thermosensitivity
CC   as compared with a parent xylanase enzyme. The variant xylanases of the
CC   invention are useful for degrading or modifying a plant cell wall, and
CC   for processing a plant material. They may be useful for baking,
CC   processing cereals, starch production, in processing wood, and enhancing
CC   the bleaching of wood pulp. They may also be useful for a variety of
CC   applications such as animal feed, flour separation (wetmilling) and paper
CC   and pulp production. In addition, they may also be useful for preparing a
CC   flour dough. The enzymes of the invention may alter or reduce the
CC   viscosity derived from the presence of hemicellulose or arabinoxylan in a
CC   solution or system comprising plant cell wall material, and for modifying
CC   food and/or feed supplement comprising xylan. Use of the variant
CC   xylanases in baking processes improves the properties of flour based
CC   doughs made from the doughs. The baked products have highly
CC   desirable characteristics with respect to blood volume, crumb structure
CC   and appearance and additionally have an extended shelf-life. The enzymes
CC   of the invention have reduced thermosensitivity and inhibitor
CC   sensitivity, which allows a reduction in the amount of xylanase required
CC   for animal feed, starch production and baking. The present sequence is
CC   the cDNA sequence which encodes the wild-type Bacillus subtilis xylanase
CC   which was used to create the variant xylanases of the invention.
XX
SQ   Sequence 645 BP; 192 A; 111 C; 163 G; 179 T; 0 U; 0 Other;

Query Match   43.9%; Score 79; DB 9; Length 645;
Best Local Similarity   79.0%; Pred. No. 3.4e-16;
Matches   94; Conservative   0; Mismatches 25; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAGAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
          ||||||||||||||||||| ||||||||||| ||||||||||||||||| |||||||||||
Db   4 ATGTTTAAGTTTAAAAGAATTCTTAGTTCTTAGTGGATTATCGGCAGCTTTAATGAGTATTAGC 63

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGACCATCACCATATCACCATATCGAGGG 119
          ||||| | |||||||||||||||||| ||| | | || | | | |||| ||
Db   64 TTGTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGG 122

RESULT 12
AAC88889
ID   AAC88889 standard; DNA; 1349 BP.
XX
AC   AAC88889;
XX
DT   05-MAR-2001 (first entry)
XX
DE   Bacillus circulans xylanase DNA sequence.
XX
KW   Bacillus circulans; xylanase; xylanase activity; XA; bleaching agent; ds.
XX
OS   Bacillus circulans.
XX
PN   WO200068396-A2.
XX
PD   16-NOV-2000.
XX
PF   12-MAY-2000; 2000WO-US013172.
XX
PR   12-MAY-1999; 99US-0133714P.
XX
PA   (XENC-) XENCOR INC.
XX

PI   Bentzien JM;
XX
DR   WPI; 2000-679800/66.
DR   P-PSDB; AAB48530.
XX
PT   Non naturally occurring XA protein with enhanced thermophilicity,
PT   alkalophilicity or thermostability relative to the naturally occurring
PT   Bacillus circulans xylanase is used in an agent for bleaching pulp.
XX
PS   Disclosure; Fig 1C; 114pp; English.
XX
CC   The present sequence is given in a specification relating to non
CC   naturally occurring xylanase activity (XA) proteins. The XA proteins
CC   comprise an amino acid sequence less than 97% identical to a naturally
CC   occurring Bacillus circulans xylanase. They are modified to exhibit
CC   enhanced thermophilicity, alkalophilicity or thermostability relative to
CC   the naturally occurring B. circulans xylanase. They may be used as the
CC   active compound in a bleaching agent which is used for bleaching pulp
XX
SQ   Sequence 1349 BP; 396 A; 232 C; 312 G; 409 T; 0 U; 0 Other;

Query Match   43.9%; Score 79; DB 3; Length 1349;
Best Local Similarity   79.0%; Pred. No. 4.5e-16;
Matches   94; Conservative   0; Mismatches 25; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAGAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
          ||||||||||||||||||| ||||||||||| ||||||||||||||||| |||||||||||
Db   392 ATGTTTAAGTTTAAAAGAATTCTTAGTTCTTAGTGGATTATCGGCAGCTTTAATGAGTATTAGC 451

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGACCATCACCATATCACCATATCGAGGG 119
          ||||| | |||||||||||||||||| ||| | | || | | | |||| ||
Db   452 TTGTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGG 510

RESULT 13
AAL48894
ID   AAL48894 standard; DNA; 1349 BP.
XX
AC   AAL48894;
XX
DT   24-OCT-2002 (first entry)
XX
DE   B circulans xylanase coding sequence.
XX
KW   Xylanase activity; enzyme; thermostable; xylanase; pulp bleaching;
KW   liquid clarification; coffee extraction; plant oil extraction; gene;
KW   starch extraction; food thickener; animal food additive; mutant; ds.
XX
OS   Bacillus circulans.
XX
PN   WO200223746-A2.
XX
PD   16-MAY-2002.
XX
PF   09-NOV-2001; 2001WO-US048018.
PR   10-NOV-2000; 2000US-00710050.
XX
PA   (XENC-) XENCOR INC.
XX
PI   Bentzien J, Dahiyat B;
XX
DR   WPI; 2002-608200/65.
XX
PT   Novel xylanase activity protein, useful in bleaching process of pulp and
PT   in food and animal feed industry, has enhanced thermostability and
PT   alkalophilicity.
XX
PS   Disclosure; Fig 1C; 121pp; English.
XX
CC   The present invention relates to a non-naturally occurring xylanase
CC   activity (XA) protein comprising an amino acid sequence less than 97%
CC   identical to a naturally occurring Bacillus circulans xylanase, where the

XX
PS Disclosure; Page 108; 112pp; English.
XX
CC The present sequence encodes an endo-beta-1,4-xylanase. The specification
CC also describes an endo-beta-1,4-xylanase inhibitor, which is obtained
CC from wheat flour. The specification also describes a mutant xylanase
CC protein. The xylanase is useful for preparing a foodstuff, preferably a
CC bakery product or a substance (e.g. a dough) for making the bakery
CC product. Wild type xylanase or mutant xylanase is useful for preparing a
CC dough that is less sticky than a dough comprising a fungal xylanase. The
CC xylanase inhibitor is useful for screening high degree resistance
CC xylanases for dough preparation. The xylanase is also useful for
CC preparing a non-sticky dough. A combination of xylanase and the inhibitor
CC is useful for calibrating and/or determining the quantity of inhibitor in
CC a wheat flour sample
XX
SQ Sequence 642 BP; 191 A; 110 C; 163 G; 178 T; 0 U; 0 Other;

Query Match 43.9%; Score 79; DB 3; Length 642;
Best Local Similarity 79.0%; Pred. No. 3.4e-16;
Matches 94; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGATTAACGGCAGCTTCATGAGTATCAGC 60
Db   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGATTATCGGCAGCTTTAATGAGTATTAGC 60

Qy 61 ATGTTTCTGCAACCGCCCTCGCAGCTGGCCCGCACCATCACCATCACCATATCGAGGG 119
Db 61 TTGTTTTCGGCAACCGCCGCCCTCGCAGCTAGCACGACTACTGGCAAAATTGGACTGATGG 119

RESULT 10
AAS13814
ID AAS13814 standard; DNA; 645 BP.
XX
AC AAS13814;
XX
DT 18-DEC-2001 (first entry)
XX
DE DNA encoding Bacillus subtilis xylanase.
XX
KW Xylanase; plant cell wall; baking; cereal; starch production; wood;
KW wood pulp bleaching; hemicellulose; arabinoxylan; foodstuff; ds.
XX
OS Bacillus subtilis.
XX
FH Key   Location/Qualifiers
FT CDS   98..645
FT   /*tag= a
FT   /product= "xylanase"
XX
PN WO200166711-A1.
XX
PD 13-SEP-2001.
XX
PF 08-MAR-2001; 2001WO-IB000426.
XX
PR 08-MAR-2000; 2000GB-00005585.
PR 27-JUN-2000; 2000GB-00015751.
XX
PA (DANI-) DANISCO AS.
XX
PI Sibbesen O, Sorensen JF;
XX
DR WPI; 2001-596834/67.
DR P-PSDB; AAU07391.
XX
PT Novel variant xylanase polypeptide or its fragment useful for degrading
PT or modifying plant cell wall, comprises amino acid modifications such
PT that the polypeptide has altered sensitivity to xylanase inhibitor.
XX
PS Example 3; Page 43; 70pp; English.
XX

CC The invention relates to a variant xylanase polypeptide (I) or its
CC fragment having xylanase activity, comprising one or more amino acid
CC modifications such that (I) or its fragment has an altered sensitivity to
CC a xylanase inhibitor as compared with the parent xylanase enzyme. (II) or
CC its coding sequence (II) is useful for degrading or modifying plant cell
CC wall or for processing a plant material by contacting the plant cell wall
CC or plant material with (I) or (II). (I) is useful for modifying plant
CC materials, and in baking, processing cereals, starch production,
CC processing wood and enhancing the bleaching of wood pulp. (I) is useful
CC for altering the viscosity derived from the presence of hemicellulose or
CC arabinoxylan in a solution or system comprising plant cell wall material.
CC (I) is useful for preparing a foodstuff such as bread, pretzels,
CC tortillas, cakes, cookies, biscuits or crackers. The present sequence
CC represents the coding sequence of Bacillus subtilis xylanase as described
CC in the method of the invention
XX
SQ Sequence 645 BP; 192 A; 111 C; 163 G; 179 T; 0 U; 0 Other;

Query Match 43.9%; Score 79; DB 5; Length 645;
Best Local Similarity 79.0%; Pred. No. 3.4e-16;
Matches 94; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGATTAACGGCAGCTTCATGAGTATCAGC 60
Db   4 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGATTATCGGCAGCTTTAATGAGTATTAGC 63

Qy 61 ATGTTTCTGCAACCGCCCTCGCAGCTGGCCCGCACCATCACCATCACCATATCGAGGG 119
Db 64 TTGTTTTCGGCAACCGCCCTCGCAGCTAGCACGACTACTGGCAAAATTGGACTGATGG 122

RESULT 11
ADC27538
ID ADC27538 standard; cDNA; 645 BP.
XX
AC ADC27538;
XX
DT 18-DEC-2003 (first entry)
XX
DE Bacillus subtilis xylanase enzyme cDNA sequence.
XX
KW xylanase; xylanase inhibitor; thermosensitivity; plant cell wall;
KW plant material; baking; processing cereal; starch production;
KW processing wood; wood pulp bleaching; animal feed; flour separation;
KW wetmilling; paper and plant dough; hemicellulose;
KW arabinoxylan; food supplement; xylan; baking process; bread volume;
KW crumb structure; crumb appearance; shelf-life; gene; ss.
XX
OS Bacillus subtilis.
XX
FH Key   Location/Qualifiers
FT CDS   4..645
FT   /*tag= b
FT   /product= "Bacillus subtilis xylanase enzyme"
FT sig_peptide   4..87
FT   /*tag= a
FT mat_peptide   88..642
FT   /*tag= c
FT   /label= Mature_Bacillus_subtilis_xylanase
XX
PN WO2003020923-A1.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-IB003797.
XX
PR 04-SEP-2001; 2001GB-00021387.
XX
PA (DANI-) DANISCO AS.
XX
PI Sibbesen O, Sorensen JF;
XX
DR WPI; 2003-332934/31.

Left column

```
CC   The present sequence encodes a mutant endo-beta-1,4-xylanase. The
CC   specification also describes an endo-beta-1,4-xylanase inhibitor, which
CC   is obtained from wheat flour. The specification also describes a mutant
CC   xylanase protein. The xylanase is useful for preparing a foodstuff,
CC   preferably a bakery product or a substance (e.g. a dough) for making the
CC   bakery product. Wild type xylanase or mutant xylanase is useful for
CC   preparing a dough that is less sticky than a dough comprising a fungal
CC   xylanase. The xylanase inhibitor is useful for screening high degree
CC   resistance xylanases for dough preparation. The xylanase is also useful
CC   for preparing a non-sticky dough. A combination of xylanase and the
CC   inhibitor is useful for calibrating and/or determining the quantity of
CC   inhibitor in a wheat flour sample
XX
SQ   Sequence 642 BP; 187 A; 115 C; 162 G; 178 T; 0 U; 0 Other;

Query Match          44.2%;  Score 79.6; DB 3;  Length 642;
Best Local Similarity  72.5%;  Pred. No. 2.1e-16;
Matches 103;  Conservative  0;  Mismatches 39;  Indels  0;  Gaps  0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACCGGCAGCTTTCATGAGTATCAGC 60
Db   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTTAATGAGTATTAGC 60
Qy  61 ATGTTTCTGCAACCGCCTCTGCAGCTGCCCGCCACCATCACCACCATATGCGAGGGA 120
Db  61 TTGTTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGGG 120
Qy 121 AGGACTCCGTTAGGTCCAGCCA 142
Db 121 GGCGGTACCGTAAACGCTGTCA 142

RESULT 8
AAA47156
ID   AAA47156 standard; DNA; 642 BP.
XX
AC   AAA47156;
XX
DT   03-OCT-2000 (first entry)
XX
DE   DNA encoding a bacterial endo-beta-1,4-xylanase mutant XM2.
XX
KW   Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
KW   dough; dough preparation; ss.
XX
OS   Synthetic.
OS   Bacillus subtilis.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..642
FT                  /*tag= a
FT                  /product= "xylanase mutant"
XX
PN   WO200039289-A2.
XX
PD   06-JUL-2000.
XX
PF   17-DEC-1999;   99WO-IB002071.
XX
PR   23-DEC-1998;   98GB-00028599.
PR   06-APR-1999;   99GB-00007805.
PR   15-APR-1999;   99GB-00008645.
XX
PA   (DANI-) DANISCO AS.
XX
PI   Sibbesen O,  Sorensen JF;
XX
DR   WPI; 2000-465744/40.
DR   P-PSDB; AAY93754.
XX
PT   Mutant xylanase protein identified using xylanase inhibitor useful for
PT   preparing non-sticky dough for bakery products.
XX
```

Right column

```
PS   Claim 3; Page 110; 112pp; English.
XX
CC   The present sequence encodes a mutant endo-beta-1,4-xylanase. The
CC   specification also describes an endo-beta-1,4-xylanase inhibitor, which
CC   specification also describes an endo-beta-1,4-xylanase inhibitor, which
CC   is obtained from wheat flour. The specification also describes a mutant
CC   xylanase protein. The xylanase is useful for preparing a foodstuff,
CC   preferably a bakery product or a substance (e.g. a dough) for making the
CC   bakery product. Wild type xylanase or mutant xylanase is useful for
CC   preparing a dough that is less sticky than a dough comprising a fungal
CC   xylanase. The xylanase inhibitor is useful for screening high degree
CC   resistance xylanases for dough preparation. The xylanase is also useful
CC   for preparing a non-sticky dough. A combination of xylanase and the
CC   inhibitor is useful for calibrating and/or determining the quantity of
CC   inhibitor in a wheat flour sample
XX
SQ   Sequence 642 BP; 189 A; 113 C; 161 G; 179 T; 0 U; 0 Other;
XX

Query Match          44.2%;  Score 79.6; DB 3;  Length 642;
Best Local Similarity  72.5%;  Pred. No. 2.1e-16;
Matches 103;  Conservative  0;  Mismatches 39;  Indels  0;  Gaps  0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACCGGCAGCTTTCATGAGTATCAGC 60
Db   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTTAATGAGTATTAGC 60
Qy  61 ATGTTTCTGCAACCGCCTCTGCAGCTGCCCGCCACCATCACCACCATATGCGAGGGA 120
Db  61 TTGTTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGGG 120
Qy 121 AGGACTCCGTTAGGTCCAGCCA 142
Db 121 GGCGGTACCGTAAACGCTGTCA 142

RESULT 9
AAA47154
ID   AAA47154 standard; DNA; 642 BP.
XX
AC   AAA47154;
XX
DT   03-OCT-2000 (first entry)
XX
DE   DNA encoding a bacterial endo-beta-1,4-xylanase protein.
XX
KW   Endo-beta-1,4-xylanase inhibitor; wheat flour; xylanase; bakery product;
KW   dough; dough preparation; ss.
XX
OS   Bacillus subtilis.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..642
FT                  /*tag= a
FT                  /product= "xylanase"
XX
PN   WO200039289-A2.
XX
PD   06-JUL-2000.
XX
PF   17-DEC-1999;   99WO-IB002071.
XX
PR   23-DEC-1998;   98GB-00028599.
PR   06-APR-1999;   99GB-00007805.
PR   15-APR-1999;   99GB-00008645.
XX
PA   (DANI-) DANISCO AS.
XX
PI   Sibbesen O,  Sorensen JF;
XX
DR   WPI; 2000-465744/40.
DR   P-PSDB; AAY93752.
XX
PT   Mutant xylanase protein identified using xylanase inhibitor useful for
PT   preparing non-sticky dough for bakery products.
```

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: October 8, 2004, 21:23:54 ; Search time 937.383 Seconds
    (without alignments)
    8322.904 Million cell updates/sec

Title: US-10-009-792C-26
Perfect score: 180
Sequence: 1 atgtttaagttaaaagaa.........tcctgctcaagtgcttagag 180

Scoring table: IDENTITY_NUC
   Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
   Maximum Match 100%
   Listing first 45 summaries

Database :   GenEmbl:*
   1: gb_ba:*
   2: gb_htg:*
   3: gb_in:*
   4: gb_om:*
   5: gb_ov:*
   6: gb_pat:*
   7: gb_ph:*
   8: gb_pl:*
   9: gb_pr:*
   10: gb_ro:*
   11: gb_sts:*
   12: gb_sy:*
   13: gb_un:*
   14: gb_vi:*
   15: em_ba:*
   16: gb_fun:*
   17: em_hum:*
   18: em_in:*
   19: em_mu:*
   20: em_om:*
   21: em_or:*
   22: em_ov:*
   23: em_pat:*
   24: em_ph:*
   25: em_pl:*
   26: em_ro:*
   27: em_sts:*
   28: em_un:*
   29: em_vi:*
   30: em_htg_hum:*
   31: em_htg_inv:*
   32: em_htg_other:*
   33: em_htg_mus:*
   34: em_htg_pln:*
   35: em_htg_rod:*
   36: em_htg_mam:*
   37: em_htg_vrt:*
   38: em_sy:*
   39: em_htgo_hum:*
   40: em_htgo_mus:*
   41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	90.8	50.4	1620	1	BSU51675	U51675 Bacillus sp
2	89.2	49.6	641	6	A36647	A36647 Sequence 2
3	89.2	49.6	1413	6	A36646	A36646 Sequence 1
4	87.6	48.7	642	6	BD234184	BD234184 Proteins.
5	87.6	48.7	642	6	AX027170	AX027170 Sequence
6	79.6	44.2	642	6	BD234186	BD234186 Proteins.
7	79.6	44.2	642	6	BD234187	BD234187 Proteins.
8	79.6	44.2	642	6	BD234188	BD234188 Proteins.
9	79.6	44.2	642	6	AX027174	AX027174 Sequence
10	79.6	44.2	642	6	AX027176	AX027176 Sequence
11	79.6	44.2	642	6	AX027178	AX027178 Sequence
12	79	43.9	642	6	BD234185	BD234185 Proteins.
13	79	43.9	642	6	AX027172	AX027172 Sequence
14	79	43.9	645	6	AX244979	AX244979 Sequence
15	79	43.9	645	6	AX717003	AX717003 Sequence
16	79	43.9	702	6	AF490979	AF490979 Bacillus
17	79	43.9	834	1	AF441773	AF441773 Bacillus
18	79	43.9	975	1	BSXYNS	X59058 Bacillus sp
19	79	43.9	1076	1	BACXYNAB	M36648 B.subtilis
20	79	43.9	1349	1	BCXYLAG	X07723 Bacillus ci
21	79	43.9	1349	6	BD247935	BD247935 Novel the
22	79	43.9	1349	6	AX047508	AX047508 Sequence
23	79	43.9	1349	6	AX608888	AX608888 Sequence
24 c	79	43.9	87500	1	AF027868	AF027868 Bacillus
25 c	79	43.9	192961	1	BSUB0011	Z99114 Bacillus su
26	75.8	42.1	702	1	AF490980	AF490980 Bacillus
27	75.8	42.1	1394	1	BSP536759	BSP536759 Bacillus
28	75.8	42.1	4558	1	BSXYLAN	BSXYLAN Bacillus su
29	72.8	40.4	657	6	AX244980	AX244980 Sequence
30	50.8	28.2	1207	6	AR000512	AR000512 Sequence
31	50.8	28.2	1207	6	E11905	E11905 DNA encodin
32	50.8	28.2	1207	6	BD017497	BD017497 Thermosta
33	50.8	28.2	1207	6	BD079063	BD079063 Bleacher
34	50.8	28.2	1207	6	BD144645	BD144645 Thermosta
35	50.6	28.1	546	6	AR202281	AR202281 Sequence
36	50.6	28.1	546	6	AR202282	AR202282 Sequence
37	50.6	28.1	546	6	AR223283	AR223283 Sequence
38	50.6	28.1	585	6	AR223284	AR223284 Sequence
39	50.6	28.1	585	6	E01423	E01423 DNA sequenc
40	50.6	28.1	585	6	I05086	I05086 Sequence 11
41	50.6	28.1	906	6	AR202204	AR202204 Sequence
42	50.6	28.1	906	6	AR223206	AR223206 Sequence
43	50.6	28.1	921	6	AR202223	AR202223 Sequence
44	50.6	28.1	921	6	AR202226	AR202226 Sequence
45	50.6	28.1	921	6	AR202229	AR202229 Sequence

ALIGNMENTS

RESULT 1
BSU51675
LOCUS   BSU51675   1620 bp   DNA   linear   BCT 31-JAN-1999
DEFINITION   Bacillus sp. endo-1,4-beta-xylanhydrolase (xynA) gene, complete cds.
ACCESSION   U51675
VERSION   U51675.1   GI:4204770
KEYWORDS   .
SOURCE   Bacillus sp.
   ORGANISM   Bacillus sp.
   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1 (bases 1 to 1620)
   AUTHORS   Kim,S.C., Jeong,K.J. and Kim,M.S.
   TITLE   Direct Submission
   JOURNAL   Submitted (18-MAR-1996) Sun Chang Kim, Biological Sciences, Korea

```
FEATURES
     source          Advanced Institute of Science and Technology, 373-1, Kusong-dong
                     Yusong-ku, Taejon 305-701, Korea
                     Location/Qualifiers
                     1..1620
                     /organism="Bacillus sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1409"
     RBS             360..365
     gene            372..1013
                     /gene="xynA"
     CDS             372..1013
                     /gene="xynA"
                     /function="xylan hydrolysis"
                     /note="similar to B. subtilis xylanase, Swiss-Prot
                     Accession Number P18429"
                     /codon_start=1
                     /transl_table=11
                     /product="endo-1,4-beta-xylanhydrolase"
                     /protein_id="AAD10834.1"
                     /db_xref="GI:4204771"
                     /translation="MFKFKKKFLVGLTAAFMSISMFSATASAAGTDYWQNWTDGGGTV
                     NAVNGSGGNYSVNWSNTGNFVVGKGWTTGSPFRTINYNAGVWAPNGNGYLTLYGWTRA
                     PLIEYYVDSWGTYRPTGTYKGTVKSDGGTYDIYTTRYNAPSIDGDNTTFQYWSVSR
                     QSKRPTGSNAAITFSNHVNAWKSHGMNLGSNWAYQVLATEGYKSSGSSNVTVW"

ORIGIN
     Query Match     50.4%;  Score 90.8;  DB 1;  Length 1620;
     Best Local Similarity 77.5%;  Pred. No. 1.3e-19;
     Matches 110;  Conservative 0;  Mismatches 32;  Indels 0;  Gaps 0;

Qy    1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  372 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC  431

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATATCGAGGGA          120
        ||||||||||||||||||||||||||||| |||||||||||||
Db  432 ATGTTTCTGCAACCGCCTCTGCAGCTGGCACAGATTACTGGCAAAATTGGACTGACGGG   491

Qy  121 AGGACTCCGTTAGGTCCAGCCA                                        142
        ||||   |||
Db  492 GGCGGGACAGTAAACGCAGTCA                                        513

RESULT 2
A36647
LOCUS       A36647        641 bp    DNA       linear   PAT 05-MAR-1997
DEFINITION  Sequence 2 from Patent EP0585617.
ACCESSION   A36647
VERSION     A36647.1  GI:2293950
KEYWORDS    .
SOURCE      Bacillus subtilis
  ORGANISM  Bacillus subtilis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1 (bases 1 to 641)
  AUTHORS   Gottschalk,M.D., Schuster,E.D. and Sproessler,B.D.
  TITLE     Bacterial Xylanase, method for its production and its application
            in manufacturing bread
  JOURNAL   Patent: EP 0585617-A 2 09-MAR-1994;
            ROEHM GMBH (DE)
COMMENT     Other publication US 5306633 940426
            Other publication FI 933519 940212
            Other publication DE 4226528 940217.
FEATURES             Location/Qualifiers
     source          1..641
                     /organism="Bacillus subtilis"
                     /mol_type="unassigned DNA"
                     /strain="RH 1321"
                     /db_xref="taxon:1423"
     CDS             1..>641
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA02246.1"
                     /db_xref="GI:2293951"
                     /db_xref="REMTREMBL:CAA02246"
                     /translation="MFKFKKKFLVGLTAAFMSISMFSATASAAGTDYWQNWTDGGGTV
                     NAVNGSGGNYSVNWSNTGNFVVGKGWTTGSPFRTINYNAGVWAPNGNGYLTLYGWTRS
                     PLIEYYVDSWGTYRPTGTYKGTVKSDGGTYDIYTTRYNAPSIDGDNTTFQYWSVR
                     QSKRPTGSNAAITFSNHVNAWKSHGMNLGSNWAYQVLATEGYKSSGSSNVTVW"

ORIGIN
     Query Match     49.6%;  Score 89.2;  DB 6;  Length 641;
     Best Local Similarity 76.8%;  Pred. No. 3.7e-19;
     Matches 109;  Conservative 0;  Mismatches 33;  Indels 0;  Gaps 0;

Qy    1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC   60

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATATCGAGGGA          120
        |||||| ||||||||||||||||||||||
Db   61 ATGTTTCGGCAACCGCCTCTGCAGCTGGCACAGATTACTGGCAAAATTGGACTGACGGG   120

Qy  121 AGGACTCCGTTAGGTCCAGCCA                                        142
        ||||   |||
Db  121 GGCGGGACAGTAAACGCAGTCA                                        142

RESULT 3
A36646
LOCUS       A36646       1413 bp    DNA       linear   PAT 05-MAR-1997
DEFINITION  Sequence 1 from Patent EP0585617.
ACCESSION   A36646
VERSION     A36646.1  GI:2293949
KEYWORDS    .
SOURCE      Bacillus subtilis
  ORGANISM  Bacillus subtilis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1 (bases 1 to 1413)
  AUTHORS   Gottschalk,M.D., Schuster,E.D. and Sproessler,B.D.
  TITLE     Bacterial Xylanase, method for its production and its application
            in manufacturing bread
  JOURNAL   Patent: EP 0585617-A 1 09-MAR-1994;
            ROEHM GMBH (DE)
COMMENT     Other publication US 5306633 940426
            Other publication FI 933519 940212
            Other publication DE 4226528 940217.
FEATURES             Location/Qualifiers
     source          1..1413
                     /organism="Bacillus subtilis"
                     /mol_type="unassigned DNA"
                     /strain="RH 1321"
                     /db_xref="taxon:1423"

ORIGIN
     Query Match     49.6%;  Score 89.2;  DB 6;  Length 1413;
     Best Local Similarity 76.8%;  Pred. No. 4.3e-19;
     Matches 109;  Conservative 0;  Mismatches 33;  Indels 0;  Gaps 0;

Qy    1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  506 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC  565

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATATCGAGGGA          120
        |||||| ||||||||||||||||||||||
Db  566 ATGTTTCGGCAACCGCCTCTGCAGCTGGCACAGATTACTGGCAAAATTGGACTGACGGG   625

Qy  121 AGGACTCCGTTAGGTCCAGCCA                                        142
        ||||   |||
Db  626 GGCGGGACAGTAAACGCAGTCA                                        647

RESULT 4
BD234184
LOCUS       BD234184      642 bp    DNA       linear   PAT 17-JUL-2003
DEFINITION  Proteins.
```

```
ACCESSION   BD234184
VERSION     BD234184.1  GI:33043954
KEYWORDS    JP 2002533121-A/1.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Sibbesen,O. and Soerensen,J.F.
  TITLE     Proteins
  JOURNAL   Patent: JP 2002533121-A 1 08-OCT-2002;
            DANISCO AS
COMMENT     OS   Artificial Sequence
            PN   JP 2002533121-A/1
            PD   08-OCT-2002
            PF   17-DEC-1999 JP 2000591181
            PR   23-DEC-1998 GB 9828599.2,06-APR-1999 GB 9907805.7 PR
            15-APR-1999 GB 9908645.6
            PI   OLE SIBBESEN,JENS FRISBAEK SOERENSEN
            PC   C12N15/09,A21D2/26,C07K14/415,C12N1/15,C12N1/19,C12N1/20 PC
            ,C12N1/21,C12N5/10,
            PC   C12N9/42,C12P21/02,C12Q1/34,C12N15/00,C12N5/00 CC
            Description of Artificial Sequence:wild-type xylanase FH Key
            FT   source          1..642
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match             48.7%;  Score 87.6;  DB 6;  Length 642;
  Best Local Similarity   76.1%;  Pred. No. 1.3e-18;
  Matches 108;  Conservative 0;  Mismatches 34;  Indels 0;  Gaps 0;

Qy          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGATTGTGATTAACGGCAGCTTTCATGAGTATCAGC 60

Qy         61 ATGTTTTCTGCAACGCCTCTGCAGCTGGCCGCCGCACCATCACCATATCGAGGGA 120
              |||||| || |||||||||||||||||||||||||||| ||    ||      ||
Db         61 ATGTTTTCGGCAACGCCTCTGCAGCTGGCCGCCGCACAGATTACTGGCAAAATTGGACTGACGGG 120

Qy        121 AGGACTCCGTTAGGTCCAGCCA 142
                |||   || ||    |||
Db        121 GGCGGGACAGTAAACGCAGTCA 142

RESULT 5
AX027170
LOCUS       AX027170                 642 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 4 from Patent WO0039289.
ACCESSION   AX027170
VERSION     AX027170.1  GI:10188157
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Soerensen,J.F. and Sibbesen,O.
  TITLE     Proteins
  JOURNAL   Patent: WO 0039289-A 4 06-JUL-2000; DANISCO (DK) ; SIBBESEN OLE (DK)
            SOERENSEN JENS FRISBAEK (DK) ; DANISCO (DK) ; SIBBESEN OLE (DK)
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="wild-type xylanase"
ORIGIN

Query Match             48.7%;  Score 87.6;  DB 6;  Length 642;
```

```
  Best Local Similarity   76.1%;  Pred. No. 1.3e-18;
  Matches 108;  Conservative 0;  Mismatches 34;  Indels 0;  Gaps 0;

Qy          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGATTAACGGCAGCTTTCATGAGTATCAGC 60

Qy         61 ATGTTTTCTGCAACGCCTCTGCAGCTGGCCGCCGCACCATCACCATATCGAGGGA 120
              |||||| || ||||||||||||||||||||||||||||
Db         61 ATGTTTTCGGCAACGCCTCTGCAGCTGGCCGCCGCACAGATTACTGGCAAAATTGGACTGACGGG 120

Qy        121 AGGACTCCGTTAGGTCCAGCCA 142
                |||   || ||
Db        121 GGCGGGACAGTAAACGCAGTCA 142

RESULT 6
BD234186
LOCUS       BD234186                 642 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Proteins.
ACCESSION   BD234186
VERSION     BD234186.1  GI:33043956
KEYWORDS    JP 2002533121-A/3.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Sibbesen,O. and Soerensen,J.F.
  TITLE     Proteins
  JOURNAL   Patent: JP 2002533121-A 3 08-OCT-2002;
            DANISCO AS
COMMENT     OS   Artificial Sequence
            PN   JP 2002533121-A/3
            PD   08-OCT-2002
            PF   17-DEC-1999 JP 2000591181
            PR   23-DEC-1998 GB 9828599.2,06-APR-1999 GB 9907805.7 PR
            15-APR-1999 GB 9908645.6
            PI   OLE SIBBESEN,JENS FRISBAEK SOERENSEN
            PC   C12N15/09,A21D2/26,C07K14/415,C12N1/15,C12N1/19,C12N1/20 PC
            ,C12N1/21,C12N5/10,
            PC   C12N9/42,C12P21/02,C12Q1/34,C12N15/00,C12N5/00 CC
            Description of Artificial Sequence:mutant xylanase FH Key
            FT   source          1..642
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match             44.2%;  Score 79.6;  DB 6;  Length 642;
  Best Local Similarity   72.5%;  Pred. No. 7.2e-16;
  Matches 103;  Conservative 0;  Mismatches 39;  Indels 0;  Gaps 0;

Qy          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
              |||||||||||||||||||||||||||||||| ||| ||||||||| ||||||| ||||
Db          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGATTATCGGCAGCTTTTAATGAGTATTAGC 60

Qy         61 ATGTTTTCTGCAACGCCTCTGCAGCTGGCCGCCGCACCATCACCATATCGAGGGA 120
               |||||| || |||||||||||||||||||||||||||
Db         61 TTGTTTTCGGCAACGCCTCTGCAGCTGGCCGCCGCACAGACTACTGGACTGATGGG 120

Qy        121 AGGACTCCGTTAGGTCCAGCCA 142
                |||   || ||
Db        121 GGCGGTACCCGTAAACCTGTCA 142

RESULT 7
BD234187
LOCUS       BD234187                 642 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Proteins.
```

```
ACCESSION   BD234187
VERSION     BD234187.1  GI:33043957
KEYWORDS    JP 2002533121-A/4.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Sibbesen,O. and Soerensen,J.F.
  TITLE     Proteins
  JOURNAL   Patent: JP 2002533121-A 4 08-OCT-2002;
            DANISCO AS
COMMENT     OS   Artificial Sequence
            PN   JP 2002533121-A/4
            PD   08-OCT-2002
            PF   17-DEC-1999 JP 2000591181
            PR   23-DEC-1998 GB 9828599.2,06-APR-1999 GB   9907805.7 PR
            15-APR-1999 GB   9908645.6
            PI   OLE SIBBESEN,JENS FRISBAEK SOERENSEN
            PC   C12N15/09,A21D2/26,C07K14/415,C12N1/15,C12N1/19,C12N1/20 PC
            ,C12N1/21,C12N5/10,
            PC   C12N9/42,C12P21/02,C12Q1/34,C12N15/00,C12N15/00 CC
            Description of Artificial Sequence:mutant xylanase FH Key.
FEATURES             Location/Qualifiers
     source          1..642
                     /organism='Artificial Sequence'.
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN
Query Match        44.2%;  Score 79.6;  DB 6;  Length 642;
  Best Local Similarity  72.5%;  Pred. No. 7.2e-16;
  Matches  103;  Conservative  0;  Mismatches  39;  Indels  0;  Gaps  0;

Qy     1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCCAGC 60
           ||||||||||||||||||| ||||||| |||| |||||||||||
Db     1 ATGTTTAAGTTTAAAAAGAATTTCTTAGTTATCGGCAGCTTTAATGAGTATTAGC 60

Qy    61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCGCCGCACCATCACCATATCGAGGGA 120
           |||| ||| || || ||||
Db    61 TTGTTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGGG 120

Qy   121 AGGACTCCGTTAGGTCCAGCCA 142
           || |||
Db   121 GGCGGTACCGTAAACGCTGTCA 142


RESULT 8
BD234188
LOCUS       BD234188                     642 bp    DNA        linear    PAT 17-JUL-2003
DEFINITION  Proteins.
ACCESSION   BD234188
VERSION     BD234188.1  GI:33043958
KEYWORDS    JP 2002533121-A/5.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Sibbesen,O. and Soerensen,J.F.
  TITLE     Proteins
  JOURNAL   Patent: JP 2002533121-A 5 08-OCT-2002;
            DANISCO AS
COMMENT     OS   Artificial Sequence
            PN   JP 2002533121-A/5
            PD   08-OCT-2002
            PF   17-DEC-1999 JP 2000591181
            PR   23-DEC-1998 GB 9828599.2,06-APR-1999 GB   9907805.7 PR
            15-APR-1999 GB   9908645.6
            PI   OLE SIBBESEN,JENS FRISBAEK SOERENSEN
            PC   C12N15/09,A21D2/26,C07K14/415,C12N1/15,C12N1/19,C12N1/20 PC
            ,C12N1/21,C12N5/10,
            PC   C12N9/42,C12P21/02,C12Q1/34,C12N15/00,C12N15/00 CC
            Description of Artificial Sequence:mutant xylanase FH Key.
FEATURES             Location/Qualifiers
     source          1..642
                     /organism='Artificial Sequence'.
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN
Query Match        44.2%;  Score 79.6;  DB 6;  Length 642;
  Best Local Similarity  72.5%;  Pred. No. 7.2e-16;
  Matches  103;  Conservative  0;  Mismatches  39;  Indels  0;  Gaps  0;

Qy     1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCCAGC 60
           ||||||||||||||||||| ||||||| |||| |||||||||||
Db     1 ATGTTTAAGTTTAAAAAGAATTTCTTAGTTATCGGCAGCTTTAATGAGTATTAGC 60

Qy    61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCGCCGCACCATCACCATATCGAGGGA 120
           |||| ||| || || ||||
Db    61 TTGTTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGGG 120

Qy   121 AGGACTCCGTTAGGTCCAGCCA 142
           || |||
Db   121 GGCGGTACCGTAAACGCTGTCA 142


RESULT 10
AX027176
LOCUS       AX027176                     642 bp    DNA        linear    PAT 16-SEP-2000
DEFINITION  Sequence 10 from Patent WO0039289.
```

```
            PC   C12N9/42,C12P21/02,C12Q1/34,C12N15/00,C12N15/00,C12N5/00 CC
            Description of Artificial Sequence:mutant xylanase FH  Key
FEATURES             Location/Qualifiers
     source          1..642
                     /organism='Artificial Sequence'.
     FT     source   1..642
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN
Query Match        44.2%;  Score 79.6;  DB 6;  Length 642;
  Best Local Similarity  72.5%;  Pred. No. 7.2e-16;
  Matches  103;  Conservative  0;  Mismatches  39;  Indels  0;  Gaps  0;

Qy     1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
           |||||||||||||||||||   |||||||  |||| |||||||||||
Db     1 ATGTTTAAGTTTAAAAAGAATTTCTTAGTTATCGGCAGCTTTAATGAGTATTAGC 60

Qy    61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCGCCGCACCATCACCATATCGAGGGA 120
           |||| ||| ||  ||  ||||
Db    61 TTGTTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGGG 120

Qy   121 AGGACTCCGTTAGGTCCAGCCA 142
           ||  |||
Db   121 GGCGGTACCGTAAACGCTGTCA 142


RESULT 9
AX027174
LOCUS       AX027174                     642 bp    DNA        linear    PAT 16-SEP-2000
DEFINITION  Sequence 8 from Patent WO0039289.
ACCESSION   AX027174
VERSION     AX027174.1  GI:10188159
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Soerensen,J.F. and Sibbesen,O.
  TITLE     Proteins
  JOURNAL   Patent: WO 0039289-A 8 06-JUL-2000;
            SOERENSEN JENS FRISBAEK (DK) ; DANISCO (DK) ; SIBBESEN OLE (DK)
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="mutant xylanase"
ORIGIN
Query Match        44.2%;  Score 79.6;  DB 6;  Length 642;
  Best Local Similarity  72.5%;  Pred. No. 7.2e-16;
  Matches  103;  Conservative  0;  Mismatches  39;  Indels  0;  Gaps  0;

Qy     1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
           |||||||||||||||||||   |||||||  |||| |||||||||||
Db     1 ATGTTTAAGTTTAAAAAGAATTTCTTAGTTATCGGCAGCTTTAATGAGTATTAGC 60

Qy    61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCGCCGCACCATCACCATATCGAGGGA 120
           |||| ||| ||  ||  ||||
Db    61 TTGTTTTCGGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGGG 120

Qy   121 AGGACTCCGTTAGGTCCAGCCA 142
           ||  |||
Db   121 GGCGGTACCGTAAACGCTGTCA 142
```

```
Db   121 GGCGGTACCGTAAACGCTGTCA 142

RESULT 12
BD234185
LOCUS       BD234185                   642 bp    DNA    linear   PAT 17-JUL-2003
DEFINITION  Proteins.
ACCESSION   BD234185
VERSION     BD234185.1  GI:33043955
KEYWORDS    JP 200253121-A/2.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Sibbesen,O. and Soerensen,J.F.
  TITLE     Proteins
  JOURNAL   Patent: JP 200253121-A 2 08-OCT-2002;
            DANISCO AS
COMMENT     OS   Artificial Sequence
            PN   JP 200253121-A/2
            PD   08-OCT-2002
            PF   17-DEC-1999 JP 2000591181
            PR   23-DEC-1998 GB   9828599.2, 06-APR-1999 GB   9907805.7 PR
            15-APR-1999 GB   9908645.6
            PI   OLE SIBBESEN,JENS FRISBAEK SOERENSEN
            PC   C12N15/09,A21D2/26,C07K14/415,C12N1/15,C12N1/19,C12N1/20 PC
            ,C12N1/21,C12N5/10,
            PC   C12N9/42,C12P21/02,C12Q1/34,C12Q1/02,C12N15/00,C12N5/00 CC
            Description of Artificial Sequence:wild-type xylanase FH   Key
            FT   source          1..642
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match             43.9%;  Score 79;  DB 6;  Length 642;
Best Local Similarity   79.0%;  Pred. No. 1.2e-15;
Matches    94;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
              ||||||||||||||||||||||||||||| ||||||| ||||||||| ||||||||| |||
Db          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTTAATGAGTATTAGC 60

Qy         61 ATGTTTTCGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATATCGAGGG 119
              |||||||||||||||||||||||||| || || ||  |||   |     || 
Db         61 TTGTTTTCGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGG 119

RESULT 13
AX027172
LOCUS       AX027172                   642 bp    DNA    linear   PAT 16-SEP-2000
DEFINITION  Sequence 6 from Patent WO0039289.
ACCESSION   AX027172
VERSION     AX027172.1  GI:10188158
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Soerensen,J.F. and Sibbesen,O.
  TITLE     Proteins
  JOURNAL   Patent: WO 0039289-A 6 06-JUL-2000;  DANISCO (DK) ; SIBBESEN OLE (DK)
            SOERENSEN JENS FRISBAEK (DK) ; DANISCO (DK)
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="wild-type xylanase"
```

```
ACCESSION   AX027176
VERSION     AX027176.1  GI:10188160
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Soerensen,J.F. and Sibbesen,O.
  TITLE     Proteins
  JOURNAL   Patent: WO 0039289-A 10 06-JUL-2000;  DANISCO (DK) ; SIBBESEN OLE (DK)
            SOERENSEN JENS FRISBAEK (DK) ; DANISCO (DK)
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="mutant xylanase"
ORIGIN

Query Match             44.2%;  Score 79.6;  DB 6;  Length 642;
Best Local Similarity   72.5%;  Pred. No. 7.2e-16;
Matches   103;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
              ||||||||||||||||||||||||||||| ||||||| ||||||||| ||||||||| |||
Db          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTTAATGAGTATTAGC 60

Qy         61 ATGTTTTCGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATATCGAGGGA 120
              |||||||||||||||||||||||||| || || ||  |||   |     || |
Db         61 TTGTTTTCGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGGG 120

Qy        121 AGGACTCCGTTAGGTCCAGCCA 142
              ||    |  ||   |   ||
Db        121 GGCGGTACCGTAAACGCTGTCA 142

RESULT 11
AX027178
LOCUS       AX027178                   642 bp    DNA    linear   PAT 16-SEP-2000
DEFINITION  Sequence 12 from Patent WO0039289.
ACCESSION   AX027178
VERSION     AX027178.1  GI:10188161
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Soerensen,J.F. and Sibbesen,O.
  TITLE     Proteins
  JOURNAL   Patent: WO 0039289-A 12 06-JUL-2000;  DANISCO (DK) ; SIBBESEN OLE (DK)
            SOERENSEN JENS FRISBAEK (DK) ; DANISCO (DK)
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="mutant xylanase"
ORIGIN

Query Match             44.2%;  Score 79.6;  DB 6;  Length 642;
Best Local Similarity   72.5%;  Pred. No. 7.2e-16;
Matches   103;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
              ||||||||||||||||||||||||||||| ||||||| ||||||||| ||||||||| |||
Db          1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTTAATGAGTATTAGC 60

Qy         61 ATGTTTTCGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATATCGAGGGA 120
              |||||||||||||||||||||||||| || || ||  |||   |     || |
Db         61 TTGTTTTCGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGGG 120

Qy        121 AGGACTCCGTTAGGTCCAGCCA 142
              ||    |  ||   |   ||
```

ORIGIN

Query Match   43.9%;   Score 79; DB 6; Length 642;
Best Local Similarity 79.0%;   Pred. No. 1.2e-15;
Matches   94; Conservative   0; Mismatches 25; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
        |||||||||||||||||||||||||||||||| ||||||||||||| ||||| |||||||
Db   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTAATGAGTATTAGC 60

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATCACCATATCGAGGG 119
        ||||||||||||||||||||||||| || ||| | ||
Db   61 TTGTTTTCGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGG 119

RESULT 14
AX244979
LOCUS   AX244979   645 bp   DNA   linear   PAT 28-SEP-2001
DEFINITION   Sequence 10 from Patent WO0166711.
ACCESSION   AX244979
VERSION   AX244979.1 GI:15859704
KEYWORDS
SOURCE   Bacillus subtilis
ORGANISM   Bacillus subtilis
             Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1
AUTHORS   Sibbesen,O. and Sorensen,J.F.
TITLE   Xylanase variants having altered sensitivity to xylanase inhibitors
JOURNAL   Patent: WO 0166711-A 10 13-SEP-2001;
             DANISCO A/S (DK)
FEATURES   Location/Qualifiers
     source   1..645
                     /organism="Bacillus subtilis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1423"

ORIGIN

Query Match   43.9%;   Score 79; DB 6; Length 645;
Best Local Similarity 79.0%;   Pred. No. 1.2e-15;
Matches   94; Conservative   0; Mismatches 25; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
        |||||||||||||||||||||||||||||||| ||||||||||||| ||||| |||||||
Db   4 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTAATGAGTATTAGC 63

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATCACCATATCGAGGG 119
        ||||||||||||||||||||||||| || ||| | ||
Db   64 TTGTTTTCGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGG 122

RESULT 15
AX717003
LOCUS   AX717003   645 bp   DNA   linear   PAT 15-APR-2003
DEFINITION   Sequence 9 from Patent WO03020923.
ACCESSION   AX717003
VERSION   AX717003.1 GI:29890250
KEYWORDS
SOURCE   Bacillus subtilis
ORGANISM   Bacillus subtilis
             Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1
AUTHORS   Sibbesen,O. and Soerensen,J.F.
TITLE   Xylanase variants
JOURNAL   Patent: WO 03020923-A 9 13-MAR-2003;
             DANISCO A/S (DK)
FEATURES   Location/Qualifiers
     source   1..645
                     /organism="Bacillus subtilis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1423"

ORIGIN

Query Match   43.9%;   Score 79; DB 6; Length 645;

Best Local Similarity 79.0%;   Pred. No. 1.2e-15;
Matches   94; Conservative   0; Mismatches 25; Indels   0; Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
        |||||||||||||||||||||||||||||||| ||||||||||||| ||||| |||||||
Db   4 ATGTTTAAGTTTAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTAATGAGTATTAGC 63

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCCGCACCATCACCATCACCATATCGAGGG 119
        ||||||||||||||||||||||||| || ||| | ||
Db   64 TTGTTTTCGCAACCGCCTCTGCAGCTAGCACAGACTACTGGCAAAATTGGACTGATGG 122

Search completed: October 9, 2004, 03:08:35
Job time : 941.383 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2004, 00:53:40 ; Search time 3228.76 Seconds
(without alignments)
4837.123 Million cell updates/sec

Title: US-10-009-792C-18_COPY_88_610
Perfect score: 523
Sequence: 1 acccccctgggcctgccag.........acgcgcacttgccagccct 523

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1493109027b residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	523	100.0	906	14	CA487627	AGENCOURT
2	523	100.0	1080	12	BM553432	AGENCOURT
3	523	100.0	1123	12	BM554035	AGENCOURT
4	522	99.8	983	12	BM557421	AGENCOURT
5	521.4	99.7	897	14	CD013924	90138839
6	521.4	99.7	1096	12	BM923410	AGENCOURT
7	521.4	99.7	1141	12	BM906188	AGENCOURT
8	506.2	96.8	1445	12	BM906367	AGENCOURT
9	504	96.4	624	12	AY414900	Homo sapi
10	499.4	95.5	993	12	BM915836	AGENCOURT
11	497.6	95.1	608	29	AY414901	Pan trogl
12	489.4	93.6	990	12	BM423896	AGENCOURT
13	485.6	92.8	986	14	CA487516	AGENCOURT
14	483	92.4	1012	12	BM556479	AGENCOURT
15	426	81.5	554	12	CB126975	K-EST0176
16	423.4	81.0	817	12	BM009358	603629823
17	419.4	80.2	874	14	CA489884	AGENCOURT
18	375.6	71.8	570	12	BI961242	MONO1_7_G
19	355.6	68.0	548	12	BI960812	MONO1_1_B
20	332.6	63.6	973	12	BI411128	602962472
21	313.8	60.0	598	12	BG548320	602575289
22	313.8	60.0	789	14	CD013925	90138939
23	313.6	60.0	642	29	AY414902	Mus muscu
24	310	59.3	829	12	BM009247	603629663
25	309.4	59.2	622	14	CD013926	90138855
26	307.8	58.9	952	14	CA487806	AGENCOURT
27	300.6	57.5	507	10	BE485194	172130 BA
28	284	54.3	346	14	CB705562	AMGNNUC:T
29	275.4	52.7	579	14	CB127794	K-EST0177
30	263.4	50.4	1375	12	BM554037	AGENCOURT
31	245.8	47.0	802	14	CA489740	AGENCOURT
32	245.8	47.0	948	12	BI822673	60303600б
33	213	40.7	368	12	BI961002	MONO1_3_H
34	200.4	38.3	466	14	CB714821	AMGNNUC:H
35	189	36.1	314	14	CB136760	K-EST0189
36	180.4	34.5	509	12	BM256261	518592 MA
37	137.4	26.3	392	13	BY022511	BY022511
38	135.8	26.0	327	10	BF848766	QVO-EN010
39	110.8	21.2	385	13	BY024867	BY024867
40	99.6	19.0	380	13	BY027389	BY027389
41	99	18.9	1123	12	BM807686	AGENCOURT
42	96.4	18.4	380	13	BY026163	BY026163
43	96.2	18.4	582	29	CE128466	tigr-gss-
44	91.8	17.6	382	13	BY026876	BY026876
45	90.2	17.2	370	13	BY023671	BY023671

ALIGNMENTS

RESULT 1
CA487627
LOCUS   CA487627   906 bp   mRNA   linear   EST 14-NOV-2002
DEFINITION AGENCOURT_10808790 MAPcL Homo sapiens cDNA clone IMAGE:6719127 5',
            mRNA sequence.
ACCESSION   CA487627
VERSION   CA487627.1 GI:24947489
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 906)
  AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE   Unpublished (1999)
  JOURNAL   NIH-MGC http://mgc.nci.nih.gov/.
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Kristi A. Egland, Ira Pastan
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14276 row: i column: 15
            High quality sequence stop: 712.

REFERENCE   1 (bases 1 to 1080)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: DCTD/DTP
               cDNA Library Preparation: Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Agencourt Bioscience Corporation
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLCM1968 row: p column: 03
               High quality sequence stop: 746.

FEATURES   Location/Qualifiers
   source   1..1080
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5467250"
               /tissue_type="amelanotic melanoma, cell line"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH_MGC_41"
               /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
               EcoRI; cDNA made by oligo-dT priming. Directionally cloned
               into EcoRI/XhoI sites using the following 5' adaptor:
               GGCACAGG(G). Library constructed by Ling Hong in the
               laboratory of Gerald M. Rubin (University of California,
               Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
               Superscript II RT (Life Technologies). Note: this is a
               NIH_MGC Library."

ORIGIN

Query Match   100.0%;   Score 523;   DB 12;   Length 1080;
Best Local Similarity 100.0%;   Pred. No. 7.9e-97;
Matches 523; Conservative   0; Mismatches   0;   Indels 0; Gaps 0;

Qy   1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 204

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCGCAGCGGCGTCCAGGAGAGCTGTGTGCCACCTACAAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 GTGAGGAAGATCCAGGGCGATGGGCGCAGCGGCGTCCAGGAGAGCTGTGTGCCACCTACAAG 264

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 CTGTGCCACCCCGAGGAGCTGGTGCTGTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 324

Qy   181 CTGAGCAGCTGCCCAGCCAGGCCCTGCCAGCTGGCAGGCCAACTCCATAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 CTGAGCAGCTGCCCAGCCAGGCCCTGCCAGCTGGCAGGCCAACTCCATAGC 384

Qy   241 GGCCTTTCCTTACCAGGGGCTCCTGGAAGGGATCTCCCCGAGTTGGGT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   385 GGCCTTTCCTTACCAGGGGCTCCTGGAAGGGATCTCCCCGAGTTGGGT 444

Qy   301 CCCACCTTGGACACACTGGAGCGTCCGACTTGCCACCATCTGGCAGCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   445 CCCACCTTGGACACACTGGAGCGTCCGACTTGCCACCATCTGGCAGCAG 504

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGGTGCTCCTAGTTGCCTCCCATCTGCAGAGCTTC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   505 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGGTGCTCCTAGTTGCCTCCCATCTGCAGAGCTTC 564

Qy   421 GCCTCTCGTTTCCAGCGCCGGGCAGGAGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   565 GCCTCTCGTTTCCAGCGCCGGGCAGGAGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 624

Qy   481 CTGGAGGTGTCGTACCGGCGTTCTACGCGCCACCTTGCCCAGCCCT 523
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   625 CTGGAGGTGTCGTACCGGCGTTCTACGCGCCACCTTGCCCAGCCCT 667

FEATURES   Location/Qualifiers
   source   1..906
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6719127"
               /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
               hTERT-HME1, LNCaP"
               /lab_host="EMDH10B"
               /clone_lib="MAPcL"
               /note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
               Subtracted with brain, liver, lung, kidney and muscle.
               Directionally cloned. Priming method: oligo-dT. Average
               insert size: 1800 bp. Library amplification: 26,000 fold.
               Kristi A. Egland, James J. Vincent, Robert Strausberg,
               Bungkook Lee & Ira Pastan: Discovery of new breast
               cancer genes encoding membrane and secreted proteins.
               Manuscript submitted."

ORIGIN

Query Match   100.0%;   Score 523;   DB 14;   Length 906;
Best Local Similarity 100.0%;   Pred. No. 7.3e-97;
Matches 523; Conservative   0; Mismatches   0;   Indels 0; Gaps 0;

Qy   1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 139

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCGCAGCGGCGTCCAGGAGAGCTGTGTGCCACCTACAAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 GTGAGGAAGATCCAGGGCGATGGGCGCAGCGGCGTCCAGGAGAGCTGTGTGCCACCTACAAG 199

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   200 CTGTGCCACCCCGAGGAGCTGGTGCTGTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 259

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGGCAGGCCAACTCCATAGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   260 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGGCAGGCCAACTCCATAGC 319

Qy   241 GGCCTTTCCTTACCAGGGGCTCCTGGAAGGGATCTCCCCGAGTTGGGT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   320 GGCCTTTCCTTACCAGGGGCTCCTGGAAGGGATCTCCCCGAGTTGGGT 379

Qy   301 CCCACCTTGGACACACTGGAGCGTCCGACGTGCCGACTTTGCCACCATCTGGCAGCAG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   380 CCCACCTTGGACACACTGGAGCGTCCGACGTGCCGACTTTGCCACCATCTGGCAGCAG 439

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGGCCCGGCCTTC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   440 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGGCCCGGCCTTC 499

Qy   421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGTCCTCCCCATCTGCAGAGCTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   500 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGTCCTCCCCATCTGCAGAGCTTC 559

Qy   481 CTGGAGGTGTCGTACCGGCGTTCTACGCGCCACCTTGCCCAGCCCT 523
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   560 CTGGAGGTGTCGTACCGGCGTTCTACGCGCCACCTTGCCCAGCCCT 602

RESULT 2
BM553432
LOCUS   BM553432   1080 bp   mRNA   linear   EST 20-FEB-2002
DEFINITION   AGENCOURT 6572502 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467250
             5', mRNA sequence.
ACCESSION   BM553432
VERSION   BM553432.1 GI:18792155
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
Db  446 CCCACCTTGGACACACTGCAGCGTCGAGCTGGAGTCGCCGACTTGCCACCACCATCTGCGACAG 505
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 ATGGGAAGAACTGGAATGGCCCCGTGCCCTGCCCGGCAGGGGGTCCTCCATCTGCAGAGCTTC 420
        ||||||||||||||||||||||||||||||||||||          |||||||||||||||||
Db  506 ATGGGAAGAACTGGAATGGCCCCGTGCCCTGCCCAGCCCAGGGTGCCACCAGCTGCGGCCTTC 565

Qy  421 GCCTCTGCTTTCAGCGCGCGGCAGGAGGGGTCCTAGTTGCTTCCCATCTGCAGAGCTTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  566 GCCTCTGCTTTCAGCGCGCGGCAGGAGGGGTCCTCAGTTGCTTCCCATCTGCAGAGCTTC 625

Qy  481 CTGGAGGTGTCGTACCGCGTTCTACGCCCACTTGCCCAGCCCT 523
        |||||||||||||||||||||||        |||||||||||||
Db  626 CTGGAGGTGTCGTACCGCGGTTCTACGGCCACCTTGCCCAGCCCT 668
```

RESULT 4
BM557421
LOCUS   BM557421   983 bp   mRNA   linear   EST 20-FEB-2002
DEFINITION   AGENCOURT_6578800 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466747
             5', mRNA sequence.
ACCESSION   BM557421
VERSION   BM557421.1   GI:18799394
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 983)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM1967 row: k   column: 04
             High quality sequence stop: 593.

FEATURES   Location/Qualifiers
     source   1..983
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5466747"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"

ORIGIN

Query Match   99.8%;   Score 522;   DB 12;   Length 983;
Best Local Similarity   99.8%;   Pred. No. 1.2e-96;
Matches 522;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy    1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCCAGAGTTCCTGCTCAAGTGCTTAGAGCAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCCAGAGTTCCTGCTCAAGTGCTTAGAGCAA 204

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 GTGAGGAAGATCCAGGGCGATGGCGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 264
```

RESULT 3
BM54035
LOCUS   BM54035   1123 bp   mRNA   linear   EST 20-FEB-2002
DEFINITION   AGENCOURT_6580951 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5469210
             5', mRNA sequence.
ACCESSION   BM54035
VERSION   BM54035.1   GI:18799290
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1123)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM1974 row: a   column: 19
             High quality sequence stop: 753.

FEATURES   Location/Qualifiers
     source   1..1123
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5469210"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"
                     /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."

ORIGIN

Query Match   100.0%;   Score 523;   DB 12;   Length 1123;
Best Local Similarity   100.0%;   Pred. No. 8e-97;
Matches 523;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCCAGAGTTCCTGCTCAAGTGCTTAGAGCAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  146 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCCAGAGTTCCTGCTCAAGTGCTTAGAGCAA 205

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  206 GTGAGGAAGATCCAGGGCGATGGCGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 265

Qy  121 CTGTGCCACCCCGAGGAGTGGTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  266 CTGTGCCACCCCGAGGAGTGGTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 325

Qy  181 CTGAGCAGCTGCCCCCAGCCCTGCAGCTGGCAGGCTCTTGAGCGGCTCCATCAGC 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  326 CTGAGCAGCTGCCCCCAGCCCTGCAGCTGGCAGGCTCTTGAGCGGCTCCATCAGC 385

Qy  241 GGCCTTTTCCTCTACCAGGGCGCTCCTGCAGCCCCTGGAAGGGATCTCCCCCAGTTGGGT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  386 GGCCTTTTCCTCTACCAGGGCGCTCCTGCAGCCCCTGGAAGGGATCTCCCCCAGTTGGGT 445

Qy  301 CCCACCTTGGACACACTGCAGCGTCGAGCTGGAGTCGCCGACTTGCCACCACCATCTGCAGCAG 360
```

```
Qy    121 CTGTGCCACCCGAGGAGCTGGTGCGTCGGACACTCTCTGGCATCCCTGGGCTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 CTGTGCCACCCGAGGAGCTGGTGCGTCGGACACTCTCTGGCATCCCTGGGCTCCC 324

Qy    181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC 384

Qy    241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 GGCCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 444

Qy    301 CCCACCTTGGACACACTGGGAATGGCCCCTGCGCCCAGGGTGCCATGTGGCCGGCTTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 CCCACCTTGGACACACTGGGAATGGCCCCTGCGCCCAGGGTGCCATGTGGCCGGCTTC 504

Qy    361 ATGGAAGAACTGGGAATGGCCCTCCAGTTGCCTCCATCTGCAGAGCTTC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 ATGGAAGAACTGGGAATGGCCCTCCAGTTGCCTCCATCTGCAGAGCTTC 564

Qy    421 GCCTCTGCTTTCCAGCGCCCAGGAGGGGTCCTAGTTGCCTCCCATCTGCCAGCCCT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 GCCTCTGCTTTCCAGCGCCCAGGAGGAGNTCCTAGTTGCCTCCCATCTGCGAGCTTC 624

Qy    481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
          |||||||||||||||||||||||||||||||||||||||||||||
Db    625 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 667
```

```
RESULT 5
CD013924
LOCUS          CD013924          897 bp    mRNA     linear    EST 21-OCT-2003
DEFINITION     90138839 Single gene library Homo sapiens cDNA, mRNA sequence.
ACCESSION      CD013924
VERSION        CD013924.1    GI:37777454
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 897)
AUTHORS        Jin,P., Fu,G.K., Wilson,A.D., Yang,J., Chien,D., Hawkins,P.R.,
               Au-Young,J. and Stuve,L.L.
TITLE          PCR isolation and cloning of novel splice variant mRNAs from known
               drug target genes
JOURNAL        Unpublished (2003)
COMMENT        Contact: Jin, P.
               Incyte Corporation
               3160 Porter Drive, Palo Alto, CA 94304, USA
               Tel: 650 621 8639
               Fax: 650 621 8965
               Email: pjin@incyte.com.
FEATURES            Location/Qualifiers
     source         1..897
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone_lib="Single gene library"
                    /note="Vector: pDrive Cloning Vector; RT-PCR was performed
                    using gene-specific primers flanking the open-reading
                    frame. PCR products were subcloned into pDrive Cloning
                    Vector and sequenced completely using M13 forward and
                    reverse primers. Sequencing gaps were closed by
                    re-sequencig using primers flanking the gapped areas."
ORIGIN
Query Match     99.7%;  Score 521.4;  DB 14;  Length 897;
Best Local Similarity 99.8%;  Pred. No. 1.5e-96;
Matches 522;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy      1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCCAGAGAGCTTCCTGCTCAGTGCTTAGAGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    275 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCCAGAGAGCTTCCTGCTCAGTGCTTAGAGCAA 334

Qy     61 GTGAGGAGATCCAGGGGCCATGGGCCAGCGCTTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    335 GTGAGGAAGATCCAGGGGCGATGGGCCAGCGCTTCCAGGAGAAGCTGTGTGCCACCTACAAG 394

Qy    121 CTGTGCCACCCGAGGAGCTGGTGTGCTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    395 CTGTGCCACCCGAGGAGCTGGTGTGCTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 454

Qy    181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    455 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC 514

Qy    241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    515 GGCCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 574

Qy    301 CCACCTTGGACACACTGCAGCGTCGCCGACTTTGCCACCACCATCTGCGAGCAG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    575 CCACCTTGGACACACTGCAGCGTCGCCGACTTTGCCACCACCATCTGCGAGCAG 634

Qy    361 ATGGAAGAACTGGGAATGGCCCTGCCCTGGCCATGCGCCGGCCTTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    635 ATGGAAGAACTGGGAATGGCCCTGCCCTGGCCATGCGCCGGCCTTC 694

Qy    421 GCCTCTGCTTTCCAGCGCCCAGGAGGGGTCTAGTTGCCTCCCATCTGCGAGAGCTTC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    695 GCCTCTGCTTTCCAGCGCCCAGGAGGGGTCTAGTTGCCTCCCATCTGCGAGAGCTTC 754

Qy    481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
          ||||||||||||||||||||||||||||||||||||||||||||
Db    755 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 797
```

```
RESULT 6
BM923410
LOCUS          BM923410          1096 bp    mRNA     linear    EST 12-MAR-2002
DEFINITION     AGENCOURT_6625827 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5759022
               5', mRNA sequence.
ACCESSION      BM923410
VERSION        BM923410.1    GI:19373789
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1096)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Life Technologies, Inc.
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Agencourt Bioscience Corporation
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLAM12803  row: m  column: 07
               High quality sequence stop: 646.
FEATURES            Location/Qualifiers
     source         1..1096
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="IMAGE:5759022"
                    /clone="IMAGE:5759022"
                    /lab_host="DH10B"
                    /clone_lib="NIH_MGC_116"
                    /note="Organ: pooled colon, kidney, stomach; Vector:
                    pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                    source anonymous pool of 3 colons, age 26 yo male, 49 yo
```

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12363 row: n column: 11
High quality sequence start: 10
High quality sequence stop: 599.

FEATURES   Location/Qualifiers
     source   1..1141
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5590090"
                  /lab_host="DH10B"
                  /clone_lib="NIH_MGC_125"
                  /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
                  Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
                  of three ovaries, from females ranging in age from 38 to
                  49 yo. Library is oligo-dT primed and directionally cloned
                  (EcoRV site is destroyed upon cloning). Average insert
                  size 2.1 kb, insert size range 1-3.5 kb. Library is
                  normalized and enriched for full-length clones and was
                  constructed by C. Gruber (Invitrogen). Research Genetics
                  tracking code 036."

ORIGIN

Query Match   99.7%;   Score 521.4;   DB 12;   Length 1141;
Best Local Similarity   99.8%;   Pred. No. 1.7e-96;
Matches   522;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy     1 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 186

Qy    61 GTGAGGAAGATCCAGGGCGATGGGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 GTGAGGAAGATCCAGGGCGATGGGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 246

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 CTGTGCCACCCCGAGGAGCTGGTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 306

Qy   181 CTGAGCAGCTGCCCAGCCAGGGCCCTGCCTGGCAGGCCCTGGCAGCCAACTCCATAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 CTGAGCAGCTGCCCAGCCAGGGCCCTGCCTGGCAGGCCCTGGCAGCCAACTCCATAGC 366

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 300
```

Qy   241 GGCCTTTCCTCTACCAGGGGGATCTCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||
Db   367 GGCCTTTCCTCTACCAGGGGGATCTCCCCGAGTTGGGT 426

Qy   301 CCCACCTTGGACACACTGGACGTGCCCAGCTTTGCCACCATCTGGCACGAG 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 CCCACCTTGGACACACTGGACGTGCCCAGCTTTGCCACCATCTGGCACGAG 486

Qy   361 ATGGAAGAACTGGGAATGGCCCCGCCGCCCCCAGGGTGCCATGCCGGCCTTC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 ATGGAAGAACTGGGAATGGCCCCGCCGCCCCCAGGGTGCCATGCCGGCCTTC 546

Qy   421 GCCTCTGCTTTCAGCGCCGGGCAGGAGGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 GCCTCTGCTTTCAGCGCCGGGCAGGAGGGGGTCCTGGTTGCCTCCCATCTGCAGAGCTTC 606

Qy   481 CTGGAGGTGTCGTACCGGCTTCTACGCCACCTTGCCCAGCCCT 523
         |||||||||||||||||||||||||||||||||||||||||||
Db   607 CTGGAGGTGTCGTACCGGCTTCTACGCCACCTTGCCCAGCCCT 649

RESULT 8
BM906367
LOCUS   BM906367   1445 bp   mRNA   linear   EST 12-MAR-2002
DEFINITION   AGENCOURT_6620268 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5590315
           5', mRNA sequence.
ACCESSION   BM906367
VERSION   BM906367.1   GI:19356746

(Left portion — RESULT 7 section)

female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

ORIGIN

Query Match   99.7%;   Score 521.4;   DB 12;   Length 1096;
Best Local Similarity   99.8%;   Pred. No. 1.7e-96;
Matches   522;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy     1 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   129 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 188

Qy    61 GTGAGGAAGATCCAGGGCGATGGGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   189 GTGAGGAAGATCCAGGGCGATGGGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 248

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 CTGTGCCACCCCGAGGAGCTGGTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 308

Qy   181 CTGAGCAGCTGCCCAGCCAGGGCCCTGCCTGGCAGGCCCTGGCAGCCAACTCCATAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309 CTGAGCAGCTGCCCAGCCAGGGCCCTGCCTGGCAGGCCCTGGCAGCCAACTCCATAGC 368

Qy   241 GGCCTTTCCTCTACCAGGGGGATCTCCCCGAGTTGGGT 300
           ||||||||||||||||||||||||||||||||||||||
Db   369 GGCCTTTCCTCTACCAGGGGGATCTCCCCGAGTTGGGT 428

Qy   301 CCCACCTTGGACACACTGGACGTGCCCAGCTTTGCCACCATCTGGCAGCAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   429 CCCACCTTGGACACACTGGACGTGCCCAGCTTTGCCACCATCTGGCAGCAG 488

Qy   361 ATGGAAGAACTGGGAATGGCCCCGCCGCCCCCAGGGTGCCATGCCGGCCTTC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   489 ATGGAAGAACTGGGAATGGCCCCGCCGCCCCCAGGGTGCCATGCCGGCCTTC 548

Qy   421 GCCTCTGCTTTCAGCGCCGGGCAGGAGGGGGTCCTGGTTGCCTCCCATCTGCAGAGCTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   549 GCCTCTGCTTTCAGCGCCGGGCAGGAGGGGGTCCTGGTTGCCTCCCATCTGCAGAGCTTC 608

Qy   481 CTGGAGGTGTCGTACCGGCTTCTACGCCACCTTGCCCAGCCCT 523
           |||||||||||||||||||||||||||||||||||||||||||
Db   609 CTGGAGGTGTCGTACCGGCTTCTACGCCACCTTGCCCAGCCCT 651
```

RESULT 7
BM906188
LOCUS   BM906188   1141 bp   mRNA   linear   EST 12-MAR-2002
DEFINITION   AGENCOURT_6621301 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5590090
           5', mRNA sequence.
ACCESSION   BM906188
VERSION   BM906188.1   GI:19355567
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1141)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Invitrogen
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation

KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1445)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12364 row: g column: 20
            High quality sequence stop: 513.
FEATURES   Location/Qualifiers
     source   1..1445
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5590315"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_125"
                     /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
                     Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
                     of three ovaries, from females ranging in age from 38 to
                     49 yo. Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 2.1 kb, insert size range 1-3.5 kb. Library is
                     normalized and enriched for full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 036."
ORIGIN

Query Match   96.8%; Score 506.2; DB 12; Length 1445;
Best Local Similarity   99.2%; Pred. No. 2.5e-93;
Matches 519; Conservative   0; Mismatches   3; Indels   1; Gaps   1;

Qy   1 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   115 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 174

Qy   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGTCCGGGAGAAGCGTGTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   175 GTGAGGAAGATCCAGGGCGATGGCCAGCGTCCGGGAGAAGCGTGTGTGCCACCTACAAG 234

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGAGACACTCTCTGGGCATCCCTGGGCTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   235 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGAGACACTCTCTGGGCATCCCTGGGCTCC 294

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCCCTGCCAACTCCATAGC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   295 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCCCTGCCAACTCCATAGC 354

Qy   241 GGCCTTTTCCTCTACCAGGGGATCCTCCCCGGAAGGGATCTCCCCCGAGTTGGGT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 GGCCTTTTCCTCTACCAGGGGATCCTCCCCGGAAGGGATCTCCCCCGAGTTGGGT 414

Qy   301 CCCACCTTGGACACTGGGACAGTCGAGTCGGCCGCCGCCACCATCTGGCAGGCAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 CCCACCTTGGACACTGGGACAGTCGAGTCGGCCGCCGCCACCATCTGGCAGGCAG 474

Qy   361 ATGGAAGAACTGGGAATGGCCGCCCTGCCCTGCAGCCAGCCAGGGTGCCGGGCCTTC 420
           |||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   475 ATGGAAGAACTGGGAATGGCCGCCCTTTCCCTGCAGCCAGCCAGGGTGCCGGGCCTTC 534

Qy   421 GCCTCTGCTTCCAGCGCCGGCAGGAGGGGTCCAGTTCCCATCTGCAGAGCTTC 480
           ||||||||||||||||||||||||||||||

Db   535 GCCTCTGCTTCCAGCGCCGGCAGGAGGGGTCCTGCTTGCTCCCATCTGCAGAGCTTC 594

Qy   481 CTGGAGGTGCGTACCGCGTTCTACGCC-ACCTTGCCCAGCCC 522
           ||||||| |||| |||| ||||||||| ||| ||||||||||
Db   595 CTGGAGGGGTCGTAACGCGTTCTACGCCGTTCTACGCCCAGCCC 637

RESULT 9
AY414900
LOCUS   AY414900   624 bp   DNA   linear   GSS 17-DEC-2003
DEFINITION Homo sapiens CSF3 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY414900
VERSION   AY414900.1 GI:39770859
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 624)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE   2 (bases 1 to 624)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..624
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>624
                     /gene="CSF3"
                     /locus_tag="HCM5378"
ORIGIN

Query Match   96.4%; Score 504; DB 29; Length 624;
Best Local Similarity   98.3%; Pred. No. 4.8e-93;
Matches 523; Conservative   0; Mismatches   0; Indels   9; Gaps   1;

Qy   1 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 150

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGTCCAGGAGAAGCT------GTGTGCC 111
           |||||||||||||||||||||||||||||||||||||||||| |||||||
Db   151 GTGAGGAAGATCCAGGGCGATGGGCAGCGTCCAGGAGAAGCTGGTGGAGTGAGTGTGCC 210

Qy   112 ACCTACAAGCTGTGCCACCCCGAGGAGCTGGTGCTGCTCTGGGACACTCTCTGGGCATCCCC 171
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   211 ACCTACAAGCTGTGCCACCCCGAGGAGCTGGTGCTGCTCTGGGACACTCTCTGGGCATCCCC 270

Qy   172 TGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCCCTGCTTGAGCCAA 231
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   271 TGGGCTCCCTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCCCTGCTTGAGCCAA 330

Qy   232 CTCCATAGCGGCCTTTCCTCTACCAGGGGTGCCTGGAAGGGATCTCCCC 291
           |||||||||||||||| ||||||||||||||||||||||||||||||||
Db   331 CTCCATAGCGGCCTTTTCCTCTACCAGGGGTGCCTGGAAGGGATCTCCCC 390

Qy   292 GAGTTGGGTCCCACCTTGGACACACTGGACGACTGCGACGAGCTGGACGTCGCCGACTTGCCACCATC 351
           |||||||||

```
Db  355  GTGAGGAAGATCCAGGGGCGGATGGCGCAGCGGCTCCAGGAGAGCTGTGTGCCACCTACAAG  414
                                                                    |||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121  CTGTGCCACCCCGAGGAGCTGGTGCTGCTCCGGCATCCCCTGGGCTCCC  180
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  415  CTGTGCCACCCCGAGGAGCTGGTGCTGCTGCGGCATCCCCTGGGCTCCC  474

Qy  181  CTGAGCAGCTGCCCAGCCAGGCCCTGCAGCTGGCAGCGTGCTTGAGCCAACTCCATAGC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  475  CTGAGCAGCTGCCCAGCCAGGCCCTGCAGCTGGCAGCGTGCTTGAGCCAACTCCATAGC  534

Qy  241  GGCCTTTCCTCACCAGGGCCGTCTCACGCCACCTTGCCCCAGTTGGGT  300
         |||||||||||  ||||||||   ||     ||       ||||||||
Db  535  GGCCTTTCCTCTACCAGGGCCTCCAGGGCCCTGGAAGGGATCTCCCCCAGTTGGGT  594

Qy  301  CCCACCTTGGACACACTGCAGCTGGACGTGCCGCGACTTTGCCACCACCATCTCGGCAGCAG  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  595  CCCACCTTGGACACACTGCAGCTGGACGTGCCGCGACTTTGCCACCACCATCTCGGCAGCAG  654

Qy  361  ATGGAGAACTGGGAATG-CCCCTGCCCTGCAGCCCACCCAGGGTGCCATGC-GGCCT  418
         ||||||||||||||||| ||||||||||||||||||||||||||||||||| |||||
Db  655  ATGGAGAACTGGGAATGGCCCCTGCCCCTGCAGCCCACCCAGGGTGCCATGCGGGGCCT  714

Qy  419  TCGCCTCTGCTTTCCAGCGCCCGGACCAGGAGGGGTCCAGTTGCCTCCCATCTGCAGAGCT  478
         |||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db  715  TCGCCTCTGCTTTCCAGCGCCCGGACCAGGAGGGGTCCGGTTGCCTCCCATCTGCAGAGCT  774

Qy  479  TCCTGGAGGTGTGTACCGGTTCTACGCCACCTTGCCCAGCCCT  523
         |||||||||||||||||||  |||||||||||||||||||||||
Db  775  TCCTGGAGGTGTGTACCGCGTTCTACGCCACCTTGCCCAGCCCT  819
```

RESULT 11
AY414901

LOCUS   AY414901   608 bp   DNA   linear   GSS 17-DEC-2003
DEFINITION Pan troglodytes CSF3 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY414901
VERSION   AY414901.1 GI:39770860
KEYWORDS   GSS.
SOURCE   Pan troglodytes (chimpanzee)
  ORGANISM Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 608)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
  PUBMED   14671302
REFERENCE   2 (bases 1 to 608)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..608
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene   <1..>608
                     /gene="CSF3"
                     /locus_tag="HCM5378"
ORIGIN

Query Match   95.1%; Score 497.6; DB 29; Length 608;

```
Db  391  GAGTTGGGTCCCACCTTGGACACACTGCAGCTGGACGTGCCGCGACTTTGCCACCACCATC  450
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  352  TGGCAGCAGATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCAGGGTGCCATG  411
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  451  TGGCAGCAGATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCAGGGTGCCATG  510

Qy  412  CCGGGCCTTGGCTCTGCTTTCCAGCGCGCCAGGAGGGGTCTCAGTTGCCTCCCATCTG  471
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  511  CCGGGCCTTGGCTCTGCTTTCCAGCGCGCCAGGAGGGGTCTCAGTTGCCTCCCATCTG  570

Qy  472  CAGAGCTTCCTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT  523
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  571  CAGAGCTTCCTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT  622
```

RESULT 10
BM915836

LOCUS   BM915836   993 bp   mRNA   linear   EST 12-MAR-2002
DEFINITION AGENCOURT_6639697 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482095
            5', mRNA sequence.
ACCESSION   BM915836
VERSION   BM915836.1 GI:19366215
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 993)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2007 row: j column: 16
            High quality sequence start: 17
            High quality sequence stop: 711.
FEATURES   Location/Qualifiers
     source   1..993
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5482095"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"
                     /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

Query Match   95.5%; Score 499.4; DB 12; Length 993;
Best Local Similarity   99.4%;
Matches 522; Conservative   0; Mismatches   1; Indels   2; Gaps   2;

```
Qy    1  ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  295  ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  354

Qy   61  GTGAGGAAGATCCAGGGGCGGATGGCGCAGCGGCTCCAGGAGAGCTGTGTGCCACCTACAAG  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 13 alignment

```
                                      /tissue_type="amelanotic melanoma, cell line"
                                      /lab_host="DH10B (phage-resistant)"
                                      /clone_lib="NIH_MGC_41"
                                      /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
                                      EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                                      into EcoRI/XhoI sites using the following 5' adaptor:
                                      GGCACGAG(G). Library constructed by Ling Hong in the
                                      laboratory of Gerald M. Rubin (University of California,
                                      Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                                      Superscript II RT (Life Technologies). Note: this is a
                                      NIH_MGC Library."

ORIGIN

Query Match       93.6%;    Score 489.4;   DB 12;   Length 990;
Best Local Similarity   97.1%;    Pred. No. 5.8e-90;
Matches 507; Conservative 0; Mismatches 14; Indels 0; Gaps 1;

Qy      1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    271 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  330

Qy     61 GTGAGGAAGATCCAGGGCGATGGGCGCAGCGCCCAGGAGAAGCTGTGTGCCACTACAAG  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    331 GTGAGGAAGATCCAGGGCGATGGGCGCAGCGCCCAGGAGAAGCTGTGTGCCACTACAAG  390

Qy    121 CTGTGCCACCCCGAGGAGCTGTGTCGTCGGCATCCCTGGGCTCCC              180
          |||||||||||||||||||||||||||||||||||||||||||||
Db    391 CTGTGCCACCCCGAGGAGCTGTGTCGTCGGCATCCCTGGGCTCCC              450

Qy    181 CTGAGCAGCTGCCCCAGCAGGGCCCTGCCCAGCTGCAGGCTGCTTGAGCAATCCATAGC  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 CTGAGCAGCTGCCCCAGCAGGGCCCTGCCCAGCTGCAGGCTGCTTGAGCAATCCATAGC  510

Qy    241 GGCCTTTCCTCTACCAGGGGGTCCTGCAGGGCCTGGAAGGGATCTCCCCCGAGTTGGGT  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511 GGCCTTTCCTCTACCAGGGGGTCCTGCAGGGCCTGGAAGGGATCTCCCCCGAGTTGGGT  570

Qy    301 CCCACCTTGACACACTGACGTGCGCGGTGCGCGACTTTGCCACCACCATCTGCAGCAG  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571 CCCACCTTGACACACTGACGTGCGCGGTGCGCGACTTTGCCACCACCATCTGCAGCAG  630

Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCCGTGCCCANCCAGGGTGCCATGCCGGCCTTC  420
          ||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||||
Db    631 ATGGAAGAACTGGGAATGGCCCCTGCCCANCAAGGTGCCATGCCGGCCTTC         690

Qy    421 GCCTCTGCTTTCCAGCGCCGGGC-AGGAGGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTT  479
          |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db    691 GNCTCTGCTTTCCAGCGCCGGGGCAAGGAGGGGGTCCTCGGTTGCCTCCCATCTGCAGAGCTT  750

Qy    480 CCTGGAGGTGTGCTACCGGTTCTACGCCTTCACGCCCAGCC                521
          |||||||||||||||||||||||||||||||||||||||||
Db    751 CCTGGAGGTGTGCTACCGGTTCTACGCCCACTTTGCCACCC               792
```

RESULT 13
CA487516
LOCUS   CA487516   986 bp   mRNA   linear   EST 14-NOV-2002
DEFINITION AGENCOURT_10809149 MAPcL Homo sapiens cDNA clone IMAGE:6719006 5', mRNA sequence.
ACCESSION   CA487516
VERSION   CA487516.1 GI:24947095
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 986)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Kristi A. Egland, Ira Pastan

RESULT 12 alignment

```
Best Local Similarity   97.6%;    Pred. No. 9.7e-92;
Matches 519; Conservative 0; Mismatches 4; Indels 9; Gaps 1;

Qy      1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  134

Qy     61 GTGAGGAAGATCCAGGGCGATGGGCGCAGCGCC------GTGTGCC               111
          |||||||||||||||||||||||||||||||||      |||||||
Db    135 GTGAGGAAGATCCAGGGCGATGGGCGCAGCGCCCAGGAGAAGCTGGTGAGTGAGTGTGCC  194

Qy    112 ACCTACAAGCTGTGCCACCCCGAGGAGCTGTGTCGTCGGCACACTCTCTGGGCATCCCC  171
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 ACCTACAAGCTGTGCCACCCCGAGGAGCTGTGTCGTCGGCACACTCTCTGGGCATCCCC  254

Qy    172 TGGGCTCCCCTGAGCAGCTGCCCCAGCAGGGCCCTGCCCAGCTGCAGGCTGCTTGAGCAA  231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 TGGGCTCCCCTGAGCAGCTGCCCCAGCAGGGCCCTGCCCAGCTGCAGGCTGCTTGAGCAA  314

Qy    232 CTCCATAGCGGGCCTTTCCTCTACCAGGGGGTCCTGCAGGGCCTGGAAGGGATCTCCCC  291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 CTCCATAGCGGGCCTTTCCTCTACCAGGGGGTCCTGCAGGGCCTGGAAGGGATCTCCCC  374

Qy    292 GAGTTGGGTCCCACCTTGACACACTGACGTGGATGCGCGACTTTGCCACCACCATC    351
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 GAGTTGGGTCCCACCTTGACACACTGACGTGGATGCGCGACTTTGCCACCACCATC    434

Qy    352 TGGCAGCAGATGGAAGAACTGGGAATGGCCCCTGCCCGTGCCCCACCAGGGTGCCATG  411
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 TGGCAGCAGATGGAAGAACTGGGAATGGCCCCTGCCCGTGCCCCACCAGGGTGCCATG  494

Qy    412 CCGGCCTTCGCCTCTTGCTTTCCAGCGCCGGGGCAGGAGGGGGTCCTTCCCATCTG    471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 CCGGCCTTCGCCTCTTGCTTTCCAGCGCCGGGGCAGGAGGGGGTCCTTGCCCATCTG   554

Qy    472 CAGAGCTTCCTGGAGGTGTGCTACCGGTTCTACGCCCAGCCCT               523
          |||||||||||||||||||||||||||||||||||||||||||
Db    555 CAGAGCTTCCTGGAGGTGTGCTACCGGTTCTACGCCCAGCCCT               606
```

RESULT 12
BM423896
LOCUS   BM423896   990 bp   mRNA   linear   EST 29-JAN-2002
DEFINITION AGENCOURT_6399123 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5517129 5', mRNA sequence.
ACCESSION   BM423896
VERSION   BM423896.1 GI:18392108
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 990)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: DCTD/DTP
  cDNA Library Preparation: Rubin Laboratory
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLCM2018 row: n column: 10
  High quality sequence stop: 542.
FEATURES   Location/Qualifiers
source   1..990
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5517129"

```
ACCESSION   BM556479
VERSION     BM556479.1  GI:18797743
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1012)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1969 row: c  column: 13
            High quality sequence start: 19
            High quality sequence stop: 745.

FEATURES             Location/Qualifiers
     source          1..1012
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5467332"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"
                     /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."

ORIGIN

Query Match       92.4%;   Score 483;   DB 12;   Length 1012;
Best Local Similarity  97.5%;   Pred. No. 1.2e-88;
Matches  512;   Conservative  0;   Mismatches  10;   Indels  3;   Gaps  2;

Qy      1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGACTTCCTGCTCAAGTGCTTAGAGCAA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    298 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGACTTCCTGCTCAAGTGCTTAGAGCAA 357

Qy     61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    358 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 417

Qy    121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    418 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 477

Qy    181 CTGAGCAGCTGCCCCCAGCCACCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    478 CTGAGCAGCTGCCCCAGCCACCAGGCCCTGGCAGGCTGCTTGAGCAACTCCATAGC 537

Qy    241 GGCCTTTTCCTTACCAGGGGATCTCCCCGAGTTGGG 300
            ||||||||||||||||||||||||||||||||||||
Db    538 GGCCTTTTCCTTACCAGGGGATCTCCCCGAGTTGGG 597

Qy    301 CCCACCTTGGACACACTGACGTGCCGACGTGCCACCATCTGCAGCAG 360
            |||||||||||||||||||||||||||||||||||||||||||||||
Db    598 CCCACCTTGGACACACTGACGTGCCGACGTGCCACCATCTGCAGCAG 657

Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCAGCCACCAGGGTGCCATGCCGGCCTTC 420
```

```
RESULT 14
BM556479
LOCUS       BM556479        1012 bp    mRNA    linear    EST 20-FEB-2002
DEFINITION  AGENCOURT_6578168 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467332
```

```
Db    658 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 717
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCTGCAGACCTTCCTGCTCAAGTGCTTAGAGCAA 188

Qy    421 GCCTCTGCTTTCCAGCGGCCGGG--CAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCT 478
          ||||||||||||||||||||||    |||||||  ||||||||||||||||||
Qy     61 GTGAGGAAGATCAGGGGCCATGGCCAGCGGCCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Db    718 GCCTCTGCTTTCCAGCGCCGGGGGAGGAGGGGGTCCTGGTTGCCTCCCATCTGCAAGCTT 777
          ||||||||||||||||| | |||||||||||||| |||||||||||||||||||||||||
Db    189 GTGAGGAAGATCAGGGGCCATGGCCAGCGGCCTCCAGGAGAAGCTGTGTGCCACCTACAAG 248

Qy    479 TCCT-GGAGGTGTCGTACCGGTTCTACGGCCACCTTGCCCAGCCC 522
          |||| |||| ||||||||||||||||||||||||||||||||||||
Qy    121 CTGTGCCACCCCGGAGGAGCTGGTGCTGCTCGGACACTCCTGGGCATCCCTGGGCTCCC 180

Db    778 TCCTGGGAGGTTTCGTACCGGGTTCTACGGCCACCTTGCCCAGCCC 822
          |||| || |||||| |||||||||||||||||||||||||||||||
Db    249 CTGTGCCACCCCCGGAGGAGCTGGTGCTGCTCGGACACTCCTGGGCATCCCTGGGCTCCC 308

RESULT 15                                           Qy    181 CTGAGCAGCTGCCCCAGCCAGCGGCTGCTTGAGCAGCTCCAACTCCATAGC 240
CB126975                                               |||||||||||||||||||||||||||||||||||||||||||||||||||
CB126975                                           Db    309 CTGAGCAGCTGCCCCAGCCAGCCCTGAGCGGCTGCTTGAGCAGCTCCAACTCCATAGC 368
LOCUS       CB126975              554 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0176151 C1SNU17 Homo sapiens cDNA clone C1SNU17-14-A07 5',
                                                  Qy    241 GGCCTTTCCTCTACCAGGGGCTCTCGCAGGGGCCCTGAGAGGGATCTCCCCGAGTTGGGT 300
            mRNA sequence.                            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACCESSION   CB126975                              Db    369 GGCCTTTTCCTCTACCAGGGGCTCTCTCCAGCGGGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 428
VERSION     CB126975.1   GI:28089170
KEYWORDS    EST.                                  Qy    301 CCCACCTTGGACACACACTGCAGCTGGACGTCGCCGACTTTGCCACCACCATCTGGCAGCAG 360
SOURCE      Homo sapiens (human)                      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
ORGANISM    Homo sapiens                          Db    429 CCCACCTTGGACACACACTGCAGCTGGACGTCGCCGACTTTGCCACCACCATCTGGCAGCAG 488
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 554)                    Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 420
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and   Db 489 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 548
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001   Qy    421 GCCTCT 426
JOURNAL     Unpublished (2002)                        ||||||
CONTACT     Kim YS                                Db    549 GCCTCT 554
COMMENT     Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea      Search completed: October  9, 2004, 04:22:26
            Tel: +82-42-860-4470                                        Job time : 3235.76 secs
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 14  row: A  column: 07
            High quality sequence stop: 554.
FEATURES             Location/Qualifiers
     source          1..554
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="C1SNU17-14-A07"
                     /sex="F"
                     /tissue_type="Uterine"
                     /cell_type="Epithelial"
                     /cell_line="SNU-17"
                     /lab_host="Top10F'"
                     /clone_lib="C1SNU17"
                     /note="Organ: Cervix; Vector: pCNS-D2; Site 1: EcoRI;
                     Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."

ORIGIN

Query Match         81.5%;  Score 426;  DB 14;  Length 554;
Best Local Similarity  100.0%;  Pred. No. 4.3e-77;
Matches 426;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCTGCAGACCTTCCTGCTCAAGTGCTTAGAGCAA 60
```

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:  October 9, 2004, 03:08:44 ; Search time 449.349 Seconds
                                      (without alignments)
                                      5900.506 Million cell updates/sec

Title:          US-10-009-792C-18_COPY_88_610
Perfect score:  523
Sequence:       1 accccctgggccctgccag........acgccacctgccagccct 523

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters:     6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

           Published Applications_NA:*
       1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
       2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
       3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
       4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
       5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
       6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
       7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
       8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
       9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
      10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
      11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
      12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
      13:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
      14:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
      15:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
      16:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
      17:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
      18:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
      19:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Database :  Published Applications_NA:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	523	100.0	531	15	US-10-009-792A-20	Sequence 20, Appl
2	523	100.0	615	15	US-10-009-792A-8	Sequence 18, Appl
3	523	100.0	630	13	US-10-609-346-17	Sequence 17, Appl
4	523	100.0	1498	12	US-10-447-315-18	Sequence 18, Appl
5	523	100.0	2352	13	US-10-609-346-7	Sequence 7, Appli
6	522	99.8	1365	10	US-09-968-362-21	Sequence 21, Appl
7	522	99.8	1371	10	US-09-968-362-17	Sequence 17, Appl
8	522	99.8	1371	10	US-09-968-362-19	Sequence 19, Appl
9	521.4	99.7	525	13	US-10-411-037-1	Sequence 1, Appli
10	521.4	99.7	525	13	US-10-411-026-1	Sequence 1, Appli
11	521.4	99.7	525	17	US-10-410-962-1	Sequence 1, Appli
12	521.4	99.7	525	17	US-10-411-049-1	Sequence 1, Appli
13	521.4	99.7	525	17	US-10-410-930-1	Sequence 1, Appli
14	521.4	99.7	525	17	US-10-410-997-1	Sequence 1, Appli
15	521.4	99.7	525	17	US-10-411-012-1	Sequence 1, Appli
16	521.4	99.7	525	17	US-10-287-994-1	Sequence 1, Appli
17	521.4	99.7	525	17	US-10-410-913-1	Sequence 1, Appli
18	521.4	99.7	1508	14	US-10-044-090-825	Sequence 825, App
19	521.4	99.7	1703	12	US-10-447-315-22	Sequence 22, Appl
20	510.2	97.6	541	9	US-09-984-186-13	Sequence 13, Appl
21	510.2	97.6	541	15	US-10-237-667-13	Sequence 13, Appl
22	510.2	97.6	541	15	US-10-237-708-13	Sequence 13, Appl
23	510.2	97.6	541	15	US-10-237-866-13	Sequence 13, Appl
24	510.2	97.6	541	15	US-10-237-871-13	Sequence 13, Appl
25	510.2	97.6	541	15	US-10-237-624-13	Sequence 13, Appl
26	510.2	97.6	541	17	US-10-702-536-13	Sequence 13, Appl
27	510.2	97.6	541	13	US-10-702-636-13	Sequence 13, Appl
28	509.2	97.4	2455	9	US-09-984-186-15	Sequence 15, Appl
29	509.2	97.4	2455	15	US-10-237-667-15	Sequence 15, Appl
30	509.2	97.4	2455	15	US-10-237-708-15	Sequence 15, Appl
31	509.2	97.4	2455	15	US-10-237-866-15	Sequence 15, Appl
32	509.2	97.4	2455	15	US-10-237-871-15	Sequence 15, Appl
33	509.2	97.4	2455	15	US-10-237-624-15	Sequence 15, Appl
34	509.2	97.4	2455	17	US-10-702-536-15	Sequence 15, Appl
35	509.2	97.4	2455	15	US-10-702-636-15	Sequence 15, Appl
36	509.2	97.4	546	17	US-10-083-446-177	Sequence 177, App
37	503.8	96.3	921	15	US-10-083-446-72	Sequence 72, Appl
38	503.8	96.3	921	15	US-10-083-446-75	Sequence 75, Appl
39	503.8	96.3	921	15	US-10-083-446-78	Sequence 78, Appl
40	503.8	96.3	966	15	US-10-083-446-73	Sequence 73, Appl
41	503.8	96.3	966	15	US-10-083-446-77	Sequence 77, Appl
42	503.8	96.3	966	15	US-10-083-446-79	Sequence 79, Appl
43	503.8	96.3	1047	15	US-10-083-446-74	Sequence 74, Appl
44	503.8	96.3	1047	15	US-10-083-446-76	Sequence 76, Appl
45	502.8	96.1	921	15	US-10-083-446-84	Sequence 84, Appl

ALIGNMENTS

```
RESULT 1
US-10-009-792A-20
; Sequence 20, Application US/10009792A
; Publication No. US2003015049A1
; GENERAL INFORMATION:
;  APPLICANT: LEE, Sang-Yup
;  APPLICANT: JEONG, Ki-Jun
;  TITLE OF INVENTION: ESCHERICHIA COLI STRAIN SECRETING HUMAN
;  TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR (G-CSF)
;  FILE REFERENCE: HYLEE60.001APC
;  CURRENT APPLICATION NUMBER: US/10/009,792A
;  CURRENT FILING DATE: 2002-10-29
;  PRIOR APPLICATION NUMBER: PCT/KR01/00549
;  PRIOR FILING DATE: 2001-03-31
;  PRIOR APPLICATION NUMBER: KR 10-2000-0017052
;  PRIOR FILING DATE: 2000-03-31
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;  LENGTH: 531
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-009-792A-20

  Query Match            100.0%;  Score 523;  DB 15;  Length 531;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-133;
  Matches 523;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  63

Qy    61 GTGAGGAAGATCCAGGGCGATGGGCGGCGCTCCAGGAGAAGCGTGTGTGCCACCTACAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64 GTGAGGAAGATCCAGGGCGATGGGCGGCGCTCCAGGAGAAGCGTGTGTGCCACCTACAAG 123

Qy   121 CTGTGCCACCCCGAGGAGCGTCGGTGCTCTCGGACACTCTCGGGCATCCCCTGGGCTCCC 180
```

RESULT 2
US-10-009-792A-18
; Sequence 18, Application US/10009792A
; Publication No. US20030153049A1
; GENERAL INFORMATION:
; APPLICANT: LEE, Sang-Yup
; APPLICANT: JEONG, Ki-Jun
; TITLE OF INVENTION: ESCHERICHIA COLI STRAIN SECRETING HUMAN
; TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR (G-CSF)
; FILE REFERENCE: HYLEE60.001APC
; CURRENT APPLICATION NUMBER: US/10/009,792A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: PCT/KR01/00549
; PRIOR FILING DATE: 2001-03-31
; PRIOR APPLICATION NUMBER: KR 10-2000-0017052
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 18
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-009-792A-18

Query Match   100.0%; Score 523; DB 15; Length 615;
Best Local Similarity   100.0%; Pred. No. 1.6e-133;
Matches 523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   88 ACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 147

Qy   61 GTGAGGAAGATCCAGGGGCGATGGGCGCAGCCTCCAGGAGAGCTGTGTGCCACTACAAG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   148 GTGAGGAAGATCCAGGGGCGATGGGCGCAGCCTCCAGGAGAGCTGTGTGCCACTACAAG 207

Qy   121 CTGTGCCACCCGAGGAGAGCTGGTGCTGCTCGGACACTCTCTGGACACTCTCTGGGCTCCC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 CTGTGCCACCCGAGGAGAGCTGGTGCTGCTCGGACACTCTCTGGACACTCTCTGGGCTCCC 267

Qy   181 CTGAGCAGCTGCCCTGCCCAGGCCCTGCAGCTGCTTGAGCCAACTCCATAGC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 CTGAGCAGCTGCCCTGCCCAGGCCCTGCAGCTGCTTGAGCCAACTCCATAGC 327

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCTCGACGGGGATCTCCTGGAAGGGATCTGGGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 GGCCTTTTCCTCTACCAGGGGCTCTCGACGGGGATCTCCTGGAAGGGATCTGGGT 387

Db   124 CTGTGCCACCCGAGGAGAGCTGGTGCTGCTCGGACACTCTCTGGCATCCCTGGGCTCCC 183

Qy   181 CTGAGCAGCTGCCCTGCCCAGGCCCTGCAGGCCCTGCAGCTGCTTGAGCCAACTCCATAGC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 CTGAGCAGCTGCCCTGCCCAGGCCCTGCAGGCCCTGCAGCTGCTTGAGCCAACTCCATAGC 243

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCTCGACGTGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 GGCCTTTTCCTCTACCAGGGGCTCTCGACGTGGCCCTGAAGGGATCTCCCCGAGTTGGGT 303

Qy   301 CCCACCTTGGACACACTGGAATGGCCCTGCGACTTGCCACCACCATCTGCAGCAG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 CCCACCTTGGACACACTGGAATGGCCCTGCGACTTGCCACCACCATCTGCAGCAG 363

Qy   361 ATGGAAGAACTGGGAATGCCCTGCCCTGCGACGTGCCCTGCGACGTGCCCTGCCATGCCGGCCTTC 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 ATGGAAGAACTGGGAATGCCCTGCCCTGCGACGTGCCCTGCGACGTGCCCTGCCATGCCGGCCTTC 423

Qy   421 GCCTCTGCTTCAGCGCGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 GCCTCTGCTTCAGCGCGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 483

Qy   481 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
             |||||||||||||||||||||||||||||||||||||||||||
Db   484 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 526

RESULT 3
US-10-609-346-17
; Sequence 17, Application US/10609346
; Publication No. US20040063635A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Zailin
; APPLICANT: Fu, Yan
; TITLE OF INVENTION: RECOMBINANT HUMAN ALBUMIN FUSION PROTEINS WITH LONG-LASTING BIOLO
; TITLE OF INVENTION: EFFECTS
; FILE REFERENCE: ZYU-0603
; CURRENT APPLICATION NUMBER: US/10/609,346
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/392,948
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 17
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-609-346-17

Query Match   100.0%; Score 523; DB 13; Length 630;
Best Local Similarity   100.0%; Pred. No. 1.6e-133;
Matches 523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97 ACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 156

Qy   61 GTGAGGAAGATCCAGGGGCGATGGGCGCAGCCTCCAGGAGAGCTGTGTGCCACTACAG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   157 GTGAGGAAGATCCAGGGGCGATGGGCGCAGCCTCCAGGAGAGCTGTGTGCCACTACAG 216

Qy   121 CTGTGCCACCCCGAGGAGAGCTGGTGCTGCTCGGACATCTCTGGGCATCCCTGGGCTCCC 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   217 CTGTGCCACCCCGAGGAGAGCTGGTGCTGCTCGGACATCTCTGGGCATCCCTGGGCTCCC 276

Qy   181 CTGAGCAGCTGCCCTGCCCAGCGCCTGGCAGGTCTGGCAACTCCATAGC 240
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   277 CTGAGCAGCTGCCCTGCCCAGCGCCTGGCAGGTCTGGCAACTCCATAGC 336

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCTCGACGGGGATCTCCCGAGTTGGGT 300
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   337 GGCCTTTTCCTCTACCAGGGGCTCTCGACGGGGATCTCCCGAGTTGGGT 396

Qy   301 CCCACCTTGGACACACTGGAATGGCCCTGCGACTTGCCACCACCATCTGCAGCAG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   397 CCCACCTTGGACACACTGGAATGGCCCTGCGACTTGCCACCACCATCTGCAGCAG 456

Qy   361 ATGGAAGAACTGGGAATGCCCTGCGACGTGCCCTGCGACGTGCCCTGCCATGCCGGCCTTC 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   457 ATGGAAGAACTGGGAATGCCCTGCGACGTGCCCTGCGACGTGCCCTGCCATGCCGGCCTTC 516

Qy   421 GCCTCTGCTTCAGCGCGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480

Qy   481 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523

Db   508 GCCTCTGCTTCAGCGCGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 567

Db   568 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 610

```
Qy   421  GCCTCTGCTTTCCAGCGCGCGGGCGAGGAGGGTCCTAGTTGCCTCCCATCTGCGAGCTTC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   517  GCCTCTGCTTTCCAGCGCGCGGGCGAGGAGGGTCCTAGTTGCCTCCCATCTGCGAGCTTC  576

Qy   481  CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCT  523
          ||||||||||||||||||||||||||||||||||||||||||||
Db   577  CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCT  619
```

RESULT 4
US-10-447-315-18
; Sequence 18, Application US/10447315
; Publication No. US2004071687A1
; GENERAL INFORMATION:
; APPLICANT: Rafii, Shahin
; APPLICANT: Heissig, Beate
; APPLICANT: Hattori, Koichi
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Adult Stem Cell Recruitment
; FILE REFERENCE: 1676.006US1
; CURRENT APPLICATION NUMBER: US/10/447,315
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: US 60/383,658
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-447-315-18

Query Match   100.0%; Score 523; DB 12; Length 1498;
Best Local Similarity   100.0%; Pred. No. 1.7e-133;
Matches 523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1  ACCCCCCTGGGCCTGCCAGCTCCCTGCCCAGCTCCTGCTCAAGTGCTTAGAGCAA   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  ACCCCCCTGGGCCTGCCAGCTCCCTGCCCAGCTCCTGCTCAAGTGCTTAGAGCAA  180

Qy    61  GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCTGGGCATCCCCTGGGCTCCC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCTGGGCATCCCCTGGGCTCCC  240

Qy   121  CTGTGCCACCCCCGAGGAGTGCTGGGACACTCTCTCTGGGCATCCCCTGGGCTCCC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  CTGTGCCACCCCCGAGGAGTGCTGGGACACTCTCTCTGGGCATCCCCTGGGCTCCC  300

Qy   181  CTGAGCAGTGCCCCAGCCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CTGAGCAGTGCCCCAGCCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC  360

Qy   241  GGCCTTTCCTCTACCAGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GGCCTTTCCTCTACCAGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT  420

Qy   301  CCCACCTTGGACACACTGGACGTGCGACGTGCCACCATCTGGCAGCAG  360
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CCCACCTTGGACACACTGGACGTGCGACGTGCCACCATCTGGCAGCAG  480

Qy   361  ATGGAAGAACTGGAATGGCCCTGGAAGGGATCTCCCCGAGTTGGGT  420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ATGGAAGAACTGGGAATGGCCCTGCCCAGCCCGTGCCATGCCGGCCTTC  540

Qy   421  GCCTCTGCTTTCCAGCGCGCGGGCGAGGAGGGTCCTAGTTGCCTCCCATCTGCGAGCTTC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GCCTCTGCTTTCCAGCGCGCGGGCGAGGAGGGTCCTAGTTGCCTCCCATCTGCGAGCTTC  600

Qy   481  CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCT  523
          ||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCT  643
```

RESULT 5
US-10-609-346-7
; Sequence 7, Application US/10609346
; Publication No. US20040063635A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Zailin
; APPLICANT: Fu, Yan
; TITLE OF INVENTION: RECOMBINANT HUMAN ALBUMIN FUSION PROTEINS WITH LONG-LASTING BIOL[OGICAL]
; TITLE OF INVENTION: EFFECTS
; FILE REFERENCE: ZYU-0603
; CURRENT APPLICATION NUMBER: US/10/609,346
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/392,948
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2352
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA of HSA-GCSF
US-10-609-346-7

Query Match   100.0%; Score 523; DB 13; Length 2352;
Best Local Similarity   100.0%; Pred. No. 1.7e-133;
Matches 523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1  ACCCCCCTGGGCCTGCCAGCTCCCTGCCCAGCTCCTGCTCAAGTGCTTAGAGCAA     60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1828  ACCCCCCTGGGCCTGCCAGCTCCCTGCCCAGCTCCTGCTCAAGTGCTTAGAGCAA   1887

Qy    61  GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCTGGGCATCCCCTGGGCTCCC    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1888  GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCTGGGCATCCCCTGGGCTCCC   1947

Qy   121  CTGTGCCACCCCCGAGGAGTGCTGGGACACTCTCTGGGCATCCCCTGGGCTCCC    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1948  CTGTGCCACCCCCGAGGAGTGCTGGGACACTCTCTGGGCATCCCCTGGGCTCCC   2007

Qy   181  CTGAGCAGTGCCCCAGCCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2008  CTGAGCAGTGCCCCAGCCAGGCCCTGGCAGGCTGCTTGAGCCAACTCCATAGC   2067

Qy   241  GGCCTTTCCTCTACCAGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2068  GGCCTTTCCTCTACCAGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT   2127

Qy   301  CCCACCTTGGACACACTGGACGTGCGACGTGCCACCATCTGGCAGCAG    360
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2128  CCCACCTTGGACACACTGGACGTGCGACGTGCCACCATCTGGCAGCAG   2187

Qy   361  ATGGAAGAACTGGAATGGCCCTGCCCAGCGCCATGCCGGCCTTC    420
          ||||||||||||||||||||||||||||||||||||||||||||
Db  2188  ATGGAAGAACTGGGAATGGCCCTGCCCAGCCCCAGCGTGCCATGCCGGCCTTC   2247

Qy   421  GCCTCTGCTTTCCAGCGCGCGGGCGAGGAGGGTCCTAGTTGCCTCCCCATCTGCGAGCTTC    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2248  GCCTCTGCTTTCCAGCGCGCGGGCGAGGAGGGTCCTAGTTGCCTCCCCATCTGCGAGCTTC   2307

Qy   481  CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCCT  523
          |||||||||||||||||||||||||||||||||||||||||||||
Db  2308  CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCCT  2350
```

RESULT 6
US-09-968-362-21
; Sequence 21, Application US/09968362
; Publication No. US2003008267A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Lee-Hwei K
; APPLICANT: Sun, Bill

; LENGTH: 1368
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hG-CSF-L-vFc gamma2 (Figure 2A)
US-09-968-362-17

Query Match   99.8%;   Score 522;   DB 10;   Length 1368;
Best Local Similarity   100.0%;   Pred. No. 3.2e-133;
Matches   522;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 165

Qy     61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTTCAGGAGAAGCTGTGTGCCACCTACAAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTTCAGGAGAAGCTGTGTGCCACCTACAAG 225

Qy    121 CTGTGCCACCCCGAGGAGCTCTCTGGGCATCCCTGGGCTCCC 180
            ||||||||||||||||||||||||||||||||||||||||||
Db    226 CTGTGCCACCCCGAGGAGCTCTCTGGGCATCCCTGGGCTCCC 285

Qy    181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCTTGAGCCAACTCCATAGC 240
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    286 CTGAGCAGCTGCCCCAGCCAGGCCCTGCTTGAGCCAACTCCATAGC 345

Qy    241 GGCCTTTTCCTCTACCAGGGGCTCCCGAGTTGGGT 300
            |||||||||||||||||||||||||||||||||||
Db    346 GGCCTTTTCCTCTACCAGGGGCTCCCGAGTTGGGT 405

Qy    301 CCCACCTTGGACACACTCGAGCTGACGTGCCGACTTTGCCACCATCTGCAGCAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 CCCACCTTGGACACACTCGAGCTGACGTGCCGACTTTGCCACCATCTGCAGCAG 465

Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCAGCCCACCCGGTGCCATGCCGGCCTTC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCAGCCCACCCGGTGCCATGCCGGCCTTC 525

Qy    421 GCCTCTGCTTCCAGCGCGCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 GCCTCTGCTTCCAGCGCGCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 585

Qy    481 CTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCC 522
            |||||||||||||||||||||||||||||||||||||||||
Db    586 CTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCC 627
```

RESULT 8
US-09-968-362-19
; Sequence 19, Application US/09968362
; Publication No. US20030082679A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Lee-Hwei K
; APPLICANT: Sun, Bill
; APPLICANT: Sun, Cecily R
; TITLE OF INVENTION: Fc fusion proteins of human granulocyte colony-stimulating factor
; TITLE OF INVENTION: biological activities
; FILE REFERENCE: 03SUN2001
; CURRENT APPLICATION NUMBER: US/09/968,362
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hG-CSF-L-vFc gamma4 (Figure 2B)
US-09-968-362-19

Query Match   99.8%;   Score 522;   DB 10;   Length 1371;
Best Local Similarity   100.0%;   Pred. No. 3.2e-133;

; APPLICANT: Sun, Cecily R
; TITLE OF INVENTION: Fc fusion proteins of human granulocyte colony-stimulating factor
; TITLE OF INVENTION: biological activities
; FILE REFERENCE: 03SUN2001
; CURRENT APPLICATION NUMBER: US/09/968,362
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hG-CSF-L-vFc gamma1 (Figure 2C)
US-09-968-362-21

Query Match   99.8%;   Score 522;   DB 10;   Length 1365;
Best Local Similarity   100.0%;   Pred. No. 3.2e-133;
Matches   522;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 165

Qy     61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTTCAGGAGAAGCTGTGTGCCACCTACAAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTTCAGGAGAAGCTGTGTGCCACCTACAAG 225

Qy    121 CTGTGCCACCCCGAGGAGCTCTCTGGGCATCCCTGGGCTCCC 180
            ||||||||||||||||||||||||||||||||||||||||||
Db    226 CTGTGCCACCCCGAGGAGCTCTCTGGGCATCCCTGGGCTCCC 285

Qy    181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCTTGAGCCAACTCCATAGC 240
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    286 CTGAGCAGCTGCCCCAGCCAGGCCCTGCTTGAGCCAACTCCATAGC 345

Qy    241 GGCCTTTTCCTCTACCAGGGGCTCCCGAGTTGGGT 300
            |||||||||||||||||||||||||||||||||||
Db    346 GGCCTTTTCCTCTACCAGGGGCTCCCGAGTTGGGT 405

Qy    301 CCCACCTTGGACACACTCGAGCTGACGTGCCGACTTTGCCACCATCTGCAGCAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 CCCACCTTGGACACACTCGAGCTGACGTGCCGACTTTGCCACCATCTGCAGCAG 465

Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCAGCCCACCCGGTGCCATGCCGGCCTTC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCAGCCCACCCGGTGCCATGCCGGCCTTC 525

Qy    421 GCCTCTGCTTCCAGCGCGCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 GCCTCTGCTTCCAGCGCGCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 585

Qy    481 CTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCC 522
            |||||||||||||||||||||||||||||||||||||||||
Db    586 CTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCC 627
```

RESULT 7
US-09-968-362-17
; Sequence 17, Application US/09968362
; Publication No. US20030082679A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Lee-Hwei K
; APPLICANT: Sun, Bill
; APPLICANT: Sun, Cecily R
; TITLE OF INVENTION: Fc fusion proteins of human granulocyte colony-stimulating factor
; TITLE OF INVENTION: biological activities
; FILE REFERENCE: 03SUN2001
; CURRENT APPLICATION NUMBER: US/09/968,362
; CURRENT FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17

; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-411-037-1

Query Match   99.7%;   Score 521.4;   DB 13;   Length 525;
Best Local Similarity   99.8%;   Pred. No. 4.5e-133;
Matches 522;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   1 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCCACCTACAAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCCACCTACAAG 60

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGTTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGTTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGCTGCTCGGACACTCTCTGGGCTCCC 180
            |||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGCTGCTCGGACACTCTCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGCCCTGGGCAGCTGGCAGCAGGATCTCCATAGC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGCCCTGGGCAGCTGGCAGCAGGATCTCCATAGC 240

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGCCCACCAGGGGTGCCATGGCCGGCCTTC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGCCCACCAGGGGTGCCATGGCCGGCCTTC 300

Qy   301 CCACCTTGGACACACTGGACGTGGACGTGGACGTGGACGTGGACGTGGACGAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCACCTTGGACACACTGGACGTGGACGTGGACGTGGACGTGGACGTGGACGAG 360

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCCACCCAGGGTGCCATGCCGGCCTTC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCCACCCAGGGTGCCATGCCGGCCTTC 420

Qy   421 GCCTCTGCTTTCCAGCGCGCGGCCAGGAGAGGGGTCCTAGGCCACCTTGCCCAGCCCT 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCTTTCCAGCGCGCGGCCAGGAGAGGGGTCCTAGGCCACCTTGCCCAGCCCT 480

Qy   481 CTGGAGGTGCGTACCGCGGTTCTACGCCACCTTGCCCAGCCCT 523
            |||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGGAGGTGCGTACCGCGGTTCTACGCCACCTTGCCCAGCCCT 523

; Matches 522;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCCACCTACAAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCCACCTACAAG 165

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGTTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 GTGAGGAAGATCCAGGGCGATGGCGCAGCGTTCCAGGAGAAGCTGTGTGCCACCTACAAG 225

Qy   121 CTGTGCCACCCCGAGGAGCTGCTCGGACACTCTCTGGGCTCCC 180
            |||||||||||||||||||||||||||||||||||||||||||
Db   226 CTGTGCCACCCCGAGGAGCTGCTCGGACACTCTCTGGGCTCCC 285

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGCCCTGGGCAGCTGGCAGCAGGATCTCCATAGC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGCCCTGGGCAGCTGGCAGCAGGATCTCCATAGC 345

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGCCCACCAGGGGTGCCATGGCCGGCCTTC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 GGCCTTTCCTCTACCAGGGGCTCCTGCAGCCCACCAGGGGTGCCATGGCCGGCCTTC 405

Qy   301 CCACCTTGGACACACTGGACGTGGACGTGGACGTGGACGTGGACGTGGACGAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 CCACCTTGGACACACTGGACGTGGACGTGGACGTGGACGTGGACGTGGACGAG 465

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCCACCCAGGGTGCCATGCCGGCCTTC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCCACCCAGGGTGCCATGCCGGCCTTC 525

Qy   421 GCCTCTGCTTTCCAGCGCGCGGCCAGGAGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 GCCTCTGCTTTCCAGCGCGCGGCCAGGAGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 585

Qy   481 CTGGAGGTGCGTACCGCGGTTCTACGCCACCTTGCCCAGCCCC 522
            |||||||||||||||||||||||||||||||||||||||||||
Db   586 CTGGAGGTGCGTACCGCGGTTCTACGCCACCTTGCCCAGCCCC 627

; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-411-026-1

Query Match   99.7%;   Score 521.4;   DB 13; Length 525;
Best Local Similarity   99.8%;   Pred. No. 4.5e-133;
Matches   522;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTCCAGGAGAAGCGTGTGCCACCTACAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTCCAGGAGAAGCGTGTGCCACCTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGTGGTGCTCGGACACTCTCTGGGCATCCCTCTGGGCTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGTGGTGCTCGGACACTCTCTGGGCATCCCTCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGAGGCCTGGCAGGCGCTGCTTGAGCCAACTCCATAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGAGGCCTGGCAGGCGCTGCTTGAGCCAACTCCATAGC 240

Qy   241 GGCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300

Qy   301 CCCACCTTGGACACACTGACGTCGCCGACTTGCGCCACCACCATCTGCAGCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACACTGACGTCGCCGACTTGCGCCACCACCATCTGCAGCAG 360

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCAGCCCTGCAGCCCCAGGGTGCCATGCGCGGCCTTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGGAAGAACTGGGAATGGCCCCTGCAGCCCTGCAGCCCCAGGGTGCCATGCGCGGCCTTC 420

Qy   421 GCCTCTGCGTTTCCAGCGCGCGGGCAGGAGGGGTCCTACGGCCACCTTGCCCATCTGCAGAGCTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCGTTTCCAGCGCGCGGGCAGGAGGGGTCCTACGGCCACCTTGCCCATCTGCAGAGCTTC 480

Qy   481 CTGGAGGTGTCGTACCGGTTCTACGGCCACCTTGCCCAGCCCT 523
          ||||||||||||||||||||||||||||||||||||||||
Db   481 CTGGAGGTGTCGTACCGGTTCTACGGCCACCTTGCCCAGCCCT 523

RESULT 11
US-10-410-962-1
; Sequence 1, Application US/10410962
; Publication No. US20040077836A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryn
; TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND
; TITLE OF INVENTION: GLYCOCONJUGATION OF G-CSF
; FILE REFERENCE: 040853-01-5054
; CURRENT APPLICATION NUMBER: US/10/410,962
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19

; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-410-962-1

Query Match   99.7%;   Score 521.4;   DB 17; Length 525;
Best Local Similarity   99.8%;   Pred. No. 4.5e-133;
Matches   522;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTCCAGGAGAAGCGTGTGCCACCTACAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTCCAGGAGAAGCGTGTGCCACCTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGTGGTGCTCGGACACTCTCTGGGCATCCCTCTGGGCTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGTGGTGCTCGGACACTCTCTGGGCATCCCTCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGAGCGCTGCTTGAGCCAACTCCATAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGAGCGCTGCTTGAGCCAACTCCATAGC 240

Qy   241 GGCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300

Qy   301 CCCACCTTGGACACACTGACGTCGCCGACTTGCGCCACCATCTGCAGCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACACTGACGTCGCCGACTTGCGCCACCATCTGCAGCAG 360

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCAGCCCTGCAGCCCCAGGGTGCCATGCGCGGCCTTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGGAAGAACTGGGAATGGCCCCTGCAGCCCTGCAGCCCCAGGGTGCCATGCGCGGCCTTC 420

Qy   421 GCCTCTGCTTTCAGCGCGTACCGGTTCTACGGCCACCTTGCCTCCCATCTGCGAGAGCTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCTTTCAGCGCGTACCGGTTCTACGGCCACCTTGCCTCCCATCTGCGAGAGCTTC 480

Qy   481 CTGGAGGTGTCGTACCGGTTCTACGGCCACCTTGCCCAGCCCT 523
          ||||||||||||||||||||||||||||||||||||||||
Db   481 CTGGAGGTGTCGTACCGGTTCTACGGCCACCTTGCCCAGCCCT 523

RESULT 12
US-10-411-049-1
; Sequence 1, Application US/10411049
; Publication No. US2004008226A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryn
; TITLE OF INVENTION: INTERFERON ALPHA: REMODELING AND GLYCOCONJUGATION OF INTERFERON
; TITLE OF INVENTION: ALPHA

```
;   FILE REFERENCE: 040853-01-5055
;   CURRENT APPLICATION NUMBER: US/10/411,049
;   CURRENT FILING DATE:  2003-04-09
;   PRIOR APPLICATION NUMBER: US 60/328,523
;   PRIOR FILING DATE: 2001-10-10
;   PRIOR APPLICATION NUMBER: US 60/344,692
;   PRIOR FILING DATE: 2001-10-19
;   PRIOR APPLICATION NUMBER: US 60/387,292
;   PRIOR FILING DATE: 2002-06-07
;   PRIOR APPLICATION NUMBER: US 60/391,777
;   PRIOR FILING DATE: 2002-06-25
;   PRIOR APPLICATION NUMBER: US 60/396,594
;   PRIOR FILING DATE: 2002-07-17
;   PRIOR APPLICATION NUMBER: US 60/404,249
;   PRIOR FILING DATE: 2002-08-16
;   PRIOR APPLICATION NUMBER: US 60/407,527
;   PRIOR FILING DATE: 2002-08-28
;   NUMBER OF SEQ ID NOS: 75
;   SOFTWARE: PatentIn version 3.2
;   SEQ ID NO 1
;     LENGTH: 525
;     TYPE: DNA
;     ORGANISM: Homo sapiens
US-10-411-049-1

Query Match         99.7%;  Score 521.4;  DB 17;  Length 525;
Best Local Similarity   99.8%;  Pred. No. 4.5e-133;
Matches 522;  Conservative    0;  Mismatches  1;  Indels    0;  Gaps    0;

Qy    1 ACCCCCTGGGCCCTGCCAGCTCCCGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCCCCTGGGCCCTGCCAGCTCCCGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy  121 CTGTGCCACCCGAGGAGAGTTGGTGTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGTGCCACCCGAGGAGAGTTGGTGTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180

Qy  181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGCTGGCAACTCCATAGC 240
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGCTGGCAACTCCATAGC 240

Qy  241 GGCCTTTTCCTCTACCAGGGGATCTCCCGAGTTGGGT 300
        |||||||||||||||||||||||||||||||||||||
Db  241 GGCCTTTTCCTCTACCAGGGGATCTCCCGAGTTGGGT 300

Qy  301 CCCACCTTGGACACACTGCGAGCTGACGTGCCGCCACCACCATCTGCAGCAG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CCCACCTTGGACACACTGCGAGCTGACGTGCCGCCACCACCATCTGCAGCAG 360

Qy  361 ATGGAAGAACTGGGAATGCCCCTGCCAGCCCACCAGGGTGCCATGCCGGCCTTC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ATGGAAGAACTGGGAATGCCCCTGCCAGCCCACCAGGGTGCCATGCCGGCCTTC 420

Qy  421 GCCTCTGCTTTCCAGCGCGCCGGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTCGCAGAGCTTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCCTCTGCTTTCCAGCGCGCCGGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTCGCAGAGCTTC 480

Qy  481 CTGGAGGTGTCGTACCGGCGTCTACGCGCCACCTTGCCCAGCCCT 523
        |||||||||||||||||||||||||||||||||||||||||||||
Db  481 CTGGAGGTGTCGTACCGGCGTCTACGCGCCACCTTGCCCAGCCCT 523
```

```
;   APPLICANT: Zopf, David
;   APPLICANT: Bayer, Robert
;   APPLICANT: Hakes, David
;   APPLICANT: Chen, Xi
;   APPLICANT: Bowe, Caryn
;   TITLE OF INVENTION: INTERFERON BETA: REMODELING AND GLYCOCONJUGATION OF INTERFERON
;   TITLE OF INVENTION: BETA
;   FILE REFERENCE: 040853-01-5056
;   CURRENT APPLICATION NUMBER: US/10/410,930
;   CURRENT FILING DATE:  2003-04-09
;   PRIOR APPLICATION NUMBER: US 60/328,523
;   PRIOR FILING DATE: 2001-10-10
;   PRIOR APPLICATION NUMBER: US 60/344,692
;   PRIOR FILING DATE: 2001-10-19
;   PRIOR APPLICATION NUMBER: US 60/387,292
;   PRIOR FILING DATE: 2002-06-07
;   PRIOR APPLICATION NUMBER: US 60/391,777
;   PRIOR FILING DATE: 2002-06-25
;   PRIOR APPLICATION NUMBER: US 60/396,594
;   PRIOR FILING DATE: 2002-07-17
;   PRIOR APPLICATION NUMBER: US 60/404,249
;   PRIOR FILING DATE: 2002-08-16
;   PRIOR APPLICATION NUMBER: US 60/407,527
;   PRIOR FILING DATE: 2002-08-28
;   NUMBER OF SEQ ID NOS: 75
;   SOFTWARE: PatentIn version 3.2
;   SEQ ID NO 1
;     LENGTH: 525
;     TYPE: DNA
;     ORGANISM: Homo sapiens
US-10-410-930-1

Query Match         99.7%;  Score 521.4;  DB 17;  Length 525;
Best Local Similarity   99.8%;  Pred. No. 4.5e-133;
Matches 522;  Conservative    0;  Mismatches  1;  Indels    0;  Gaps    0;

Qy    1 ACCCCCTGGGCCCTGCCAGCTCCCGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCCCCTGGGCCCTGCCAGCTCCCGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy  121 CTGTGCCACCCGAGGAGAGTTGGTGTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGTGCCACCCGAGGAGAGTTGGTGTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180

Qy  181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGCTGGCAACTCCATAGC 240
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGAGCAGCTGCCCCAGCCAGGCCCTGGCAGCTGGCAACTCCATAGC 240

Qy  241 GGCCTTTTCCTCTACCAGGGGATCTCCCGAGTTGGGT 300
        |||||||||||||||||||||||||||||||||||||
Db  241 GGCCTTTTCCTCTACCAGGGGATCTCCCGAGTTGGGT 300

Qy  301 CCCACCTTGGACACACTGCGAGCTGACGTGCCGCCACCACCATCTGCAGCAG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CCCACCTTGGACACACTGCGAGCTGACGTGCCGCCACCACCATCTGCAGCAG 360

Qy  361 ATGGAAGAACTGGGAATGCCCCTGCCAGCCCATGCCATGCCGGCCTTC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ATGGAAGAACTGGGAATGCCCCTGCCAGCCCATGCCATGCCGGCCTTC 420

Qy  421 GCCTCTGCTTTCCAGCGCGCCGGGGCAGGAGGGGTCCTCCCATCTGCCTTCCCATCTGCAGAGCTTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCCTCTGCTTTCCAGCGCGCCGGGGCAGGAGGGGTCCTCCCATCTGCCTTCCCATCTGCAGAGCTTC 480

Qy  481 CTGGAGGTGTCGTACCGGTTCACGGCCACCTTGCCCCAGCCCT 523
        |||||||||||||||||||||||||||||||||||||||||
Db  481 CTGGAGGTGTCGTACCGGTTCTACGGCCACCTTGCCCCAGCCCT 523
```

```
RESULT 13
US-10-410-930-1
; Sequence 1, Application US/10410930
; Publication No. US20040115168A1
; GENERAL INFORMATION:
;   APPLICANT: Neose Technologies, Inc.
;   APPLICANT: DeFrees, Shawn
```

RESULT 14
US-10-410-997-1
; Sequence 1, Application US/10410997
; Publication No. US2004012838A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryn
; TITLE OF INVENTION: FOLLICLE STIMULATING HORMONE: REMODELING AND GLYCOCONJUGATION OF
; TITLE OF INVENTION: FSH
; FILE REFERENCE: 040853-01-5059
; CURRENT APPLICATION NUMBER: US/10/410,997
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-410-997-1

Query Match   99.7%; Score 521.4; DB 17; Length 525;
Best Local Similarity 99.8%; Pred. No. 4.5e-133;
Matches 522; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 ACCCCCCTGGCCTGCCAGCTCCCTGCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA   60
         |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCCTGGGCCTGCCAGCTCCCTGCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA   60

Qy   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTCCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTCCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy 121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCGGGCATCCCTGGGCTCCC   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCGGGCATCCCTGGGCTCCC   180

Qy 181 CTGAGCAGCTGCCCCAGCCAGGCCAGCGTGCTGCTTGAGCCAACTCCATAGC   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CTGAGCAGCTGCCCCAGCCAGGCCAGCGTGCTGCTTGAGCCAACTCCATAGC   240

Qy 241 GGCCTTTTCCTCTACCAGGGGGCTCTGCAGGGCCCTGGAAGGATCTCCCCGAGTTGGGT   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GGCCTTTTCCTCTACCAGGGGGCTCTGCAGGGCCCTGGAAGGATCTCCCCGAGTTGGGT   300

Qy 301 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGCACCACCAGGGTGCCGCGGCCTTC   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGCACCACCAGGGTGCCGCGGCCTTC   360

Qy 361 ATGGAAGAAGAACTGGGAATGGCCCGCCTGCC   420
         ||||||||||||||||||||||||||||||||
Db 361 ATGGAAGAAGAACTGGGAATGGCCCGCCTGCC   420

Qy 421 GCCTCTGCTTTCCAGCGCCGGCCGGCGGGAGGAGGGGTCCTGGTTGCTCCCATCTGCGAGAGCTTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GCCTCTGCTTTCCAGCGCCGGCCGGCGGGAGGAGGGGTCCTGGTTGCTCCCATCTGCGAGAGCTTC 480

Qy 481 CTGGAGGTGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT   523
         |||||||||||||||||||||||||||||||||||||||||
Db 481 CTGGAGGTGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT   523

RESULT 15
US-10-411-012-1
; Sequence 1, Application US/10411012
; Publication No. US20040132640A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bowe, Caryne
; TITLE OF INVENTION: GLYCOPEGYLATION METHODS AND PROTEINS/PEPTIDES PRODUCED BY THE
; TITLE OF INVENTION: METHODS
; FILE REFERENCE: 040853-01-5051
; CURRENT APPLICATION NUMBER: US/10/411,012
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-411-012-1

Query Match   99.7%; Score 521.4; DB 17; Length 525;
Best Local Similarity 99.8%; Pred. No. 4.5e-133;
Matches 522; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 ACCCCCCTGGCCTGCCAGCTCCCTGCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA   60
         |||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCCTGGGCCTGCCAGCTCCCTGCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA   60

Qy   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTCCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTCCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy 121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCGGGCATCCCTGGGCTCCC   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCGGGCATCCCTGGGCTCCC   180

Qy 181 CTGAGCAGCTGCCCCAGCCAGGCCAGCGTGCTGCTTGAGCCAACTCCATAGC   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CTGAGCAGCTGCCCCAGCCAGGCCAGCGTGCTGCTTGAGCCAACTCCATAGC   240

Qy 241 GGCCTTTTCCTCTACCAGGGGGCTCTGCAGGGCCCTGGAAGGATCTCCCCGAGTTGGGT   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GGCCTTTTCCTCTACCAGGGGGCTCTGCAGGGCCCTGGAAGGATCTCCCCGAGTTGGGT   300

Qy 301 CCCACCTTGGACACACTGCAGCTGCGCCGACTTGCCGCACCATCTGGCAGCAG   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CCCACCTTGGACACACTGCAGCTGCGCCGACTTGCCGCACCATCTGGCAGCAG   360

```
Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 420

Qy    421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
          ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db    421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGTCCTGGTTGCCTCCCATCTGCAGAGCTTC 480

Qy    481 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
          ||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
```

Search completed: October 9, 2004, 06:22:17
Job time : 452.349 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2004, 01:28:34 ; Search time 80.3471 Seconds
(without alignments)
3612.319 Million cell updates/sec

Title: US-10-009-792C-18_COPY_88_610
Perfect score: 523
Sequence: 1 acccccctgggcctgccag........acgccaccttgccagccct 523

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	523	100.0	525	1	US-08-434-411-1	Sequence 1, Appli
2	523	100.0	525	1	US-08-434-402-1	Sequence 1, Appli
3	523	100.0	525	1	US-08-783-288-1	Sequence 1, Appli
4	523	100.0	525	2	US-08-890-640-1	Sequence 1, Appli
5	523	100.0	525	6	5194592-25	Patent No. 5194592
6	510.2	97.6	541	4	US-08-797-689-13	Sequence 13, Appl
7	510.2	97.6	541	4	US-09-984-186-13	Sequence 13, Appl
8	510.2	97.6	2382	1	US-08-256-938-1	Sequence 1, Appli
9	509.2	97.4	2455	2	US-08-256-938-3	Sequence 3, Appli
10	509.2	97.4	2455	2	US-08-797-689-15	Sequence 15, Appl
11	509.2	97.4	2455	4	US-09-984-186-15	Sequence 15, Appl
12	507	96.9	525	4	US-08-149-101A-1	Sequence 1, Appli
13	507	96.9	525	5	PCT-US94-12873-1	Sequence 1, Appli
14	503.8	96.3	546	3	US-08-469-318-177	Sequence 177, App
15	503.8	96.3	546	3	US-08-468-609A-177	Sequence 177, App
16	503.8	96.3	546	4	US-08-446-872A-177	Sequence 177, App
17	503.8	96.3	546	4	US-08-762-227A-177	Sequence 177, App
18	503.8	96.3	546	5	PCT-US95-01185-177	Sequence 177, App
19	503.8	96.3	921	3	US-08-469-318-72	Sequence 72, Appl
20	503.8	96.3	921	3	US-08-469-318-75	Sequence 75, Appl
21	503.8	96.3	921	3	US-08-469-318-78	Sequence 78, Appl
22	503.8	96.3	921	3	US-08-468-609A-72	Sequence 72, Appl
23	503.8	96.3	921	3	US-08-468-609A-75	Sequence 75, Appl
24	503.8	96.3	921	4	US-08-468-609A-78	Sequence 78, Appl
25	503.8	96.3	921	4	US-08-446-872A-72	Sequence 72, Appl
26	503.8	96.3	921	4	US-08-446-872A-75	Sequence 75, Appl
27	503.8	96.3	921	4	US-08-446-872A-78	Sequence 78, Appl
28	503.8	96.3	921	4	US-08-762-227A-72	Sequence 72, Appl
29	503.8	96.3	921	4	US-08-762-227A-75	Sequence 75, Appl
30	503.8	96.3	921	4	US-08-762-227A-78	Sequence 78, Appl
31	503.8	96.3	921	5	PCT-US95-01185-72	Sequence 72, Appl
32	503.8	96.3	921	5	PCT-US95-01185-75	Sequence 75, Appl
33	503.8	96.3	921	5	PCT-US95-01185-78	Sequence 78, Appl
34	503.8	96.3	966	3	US-08-469-318-73	Sequence 73, Appl
35	503.8	96.3	966	3	US-08-469-318-77	Sequence 77, Appl
36	503.8	96.3	966	3	US-08-469-318-79	Sequence 79, Appl
37	503.8	96.3	966	3	US-08-468-609A-73	Sequence 73, Appl
38	503.8	96.3	966	3	US-08-468-609A-77	Sequence 77, Appl
39	503.8	96.3	966	3	US-08-468-609A-79	Sequence 79, Appl
40	503.8	96.3	966	4	US-08-446-872A-73	Sequence 73, Appl
41	503.8	96.3	966	4	US-08-446-872A-77	Sequence 77, Appl
42	503.8	96.3	966	4	US-08-446-872A-79	Sequence 79, Appl
43	503.8	96.3	966	4	US-08-762-227A-73	Sequence 73, Appl
44	503.8	96.3	966	4	US-08-762-227A-77	Sequence 77, Appl
45	503.8	96.3	966	4	US-08-762-227A-79	Sequence 79, Appl

ALIGNMENTS

RESULT 1
US-08-434-411-1
; Sequence 1, Application US/08434411
; Patent No. 5681720
; GENERAL INFORMATION:
; APPLICANT: KUGA, TETSURO
; APPLICANT: MIYAJI, HIROMASA
; APPLICANT: SATO, MORIYUKI
; APPLICANT: OKABE, MASAMI
; APPLICANT: MORIMOTO, MAKOTO
; APPLICANT: ITOH, SEIGA
; APPLICANT: YAMASAKI, MOTOO
; APPLICANT: YOKOO, YOSHIHARU
; APPLICANT: YAMAGUCHI, KAZUO
; APPLICANT: YOSHIDA, HAJIME
; APPLICANT: YOSHINORI, KOMATSU
; TITLE OF INVENTION: NOVEL POLYPEPTIDES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,411
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306799/86
; FILING DATE: 23-DEC-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 51357/88
; FILING DATE: 04-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 80088/88
; FILING DATE: 31-MAR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR
; REGISTRATION NUMBER: 25327
; REFERENCE/DOCKET NUMBER: 249-73
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100

; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: NIXON & VANDERHYE P.C.
;   STREET: 1100 NORTH GLEBE ROAD
;   CITY: ARLINGTON
;   STATE: VIRGINIA
;   COUNTRY: U.S.A.
;   ZIP: 22201-4714
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/434,402
;   FILING DATE: 03-MAY-1995
;   CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 306799/86
;   FILING DATE: 23-DEC-1986
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 51357/88
;   FILING DATE: 04-MAR-1988
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 80088/88
;   FILING DATE: 31-MAR-1988
; ATTORNEY/AGENT INFORMATION:
;   NAME: CRAWFORD, ARTHUR
;   REGISTRATION NUMBER: 25327
;   REFERENCE/DOCKET NUMBER: 249-72
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703)816-4000
;   TELEFAX: (703)816-4100
;   TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 525 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..522
;   OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
;   OTHER INFORMATION: STIMULATING FACTOR"
US-08-434-402-1

Query Match   100.0%; Score 523; DB 1; Length 525;
Best Local Similarity   100.0%; Pred. No. 5.5e-115;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGCCATGGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGCCATGGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy 121 CTGTGCCACCCCGAGGAGCTGGTCGGAGCTCTCTCGGAACACTCTCTGGGCATCCCCTGGGCTCCC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGTGCCACCCCGAGGAGCTGGTCGGAGCTCTCTCGGAACACTCTCTGGGCATCCCCTGGGCTCCC 180

Qy 181 CTGAGCAGCTGCCCAGCCAGCCCTGCAGGTGGCAGGCCCTGGAAGGATCCCCGAGTTGGGT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CTGAGCAGCTGCCCAGCCAGCCCTGCAGGTGGCAGGCCCTGGAAGGATCCCCGAGTTGGGT 240

Qy 241 GGCCTTTCCTCTACCAGGGGCTCGCCAGCCCTTGCCACCACCATCTCGGCCAGCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GGCCTTTCCTCTACCAGGGGCTCGCCAGCCCTTGCCACCACCATCTCGGCCAGCAG 300

Qy 301 CCCACCTTGGACACACTGCAGCTGACGTGCGCCGACTTTGCCACCACATCTCGGCCAGCAG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CCCACCTTGGACACACTGCAGCTGACGTGCGCCGACTTTGCCACCACATCTCGGCCAGCAG 360

Qy 361 ATGGAAGAACTGGAATGGCCCTGCCCTGCGAGCCACCCAGGGTGCCATGCGGCCCTTC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 ATGGAAGAACTGGAATGGCCCTGCCCTGCGAGCCACCCAGGGTGCCATGCGGCCCTTC 420

Qy 421 GCCTCTGCTTCCAGCGCCGGGCAGGAGGGTCCTAGTTGCCTCCCCATCTGCAGAGCTTC 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GCCTCTGCTTCCAGCGCCGGGCAGGAGGGTCCTCCTAGTTGCCTCCCCATCTGCAGAGCTTC 480

Qy 481 CTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT 523
        ||||||||||||||||||||||||||||||||||||||||||
Db 481 CTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT 523

RESULT 2
US-08-434-402-1
; Sequence 1, Application US/08434402
; Patent No. 5714581
; GENERAL INFORMATION:
;   APPLICANT: KUGA, TETSURO
;   APPLICANT: MIYAJI, HIROMASA
;   APPLICANT: SATO, MORIYUKI
;   APPLICANT: OKABE, MASAMI
;   APPLICANT: MORIMOTO, MAKOTO
;   APPLICANT: ITOH, SEIGA
;   APPLICANT: YAMASAKI, MOTOO
;   APPLICANT: YOKOO, YOSHIHARU
;   APPLICANT: YAMAGUCHI, KAZUO
;   APPLICANT: YOSHIDA, HAJIME
;   APPLICANT: YOSHINORI, KOMATSU
; TITLE OF INVENTION: NOVEL POLYPEPTIDES

; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 525 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..522
;   OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
;   OTHER INFORMATION: STIMULATING FACTOR"
US-08-434-411-1

Query Match   100.0%; Score 523; DB 1; Length 525;
Best Local Similarity   100.0%; Pred. No. 5.5e-115;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGCCATGGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGCCATGGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy 121 CTGTGCCACCCCGAGGAGCTGGTCGGAGCTCTCTCGGAACACTCTCTGGGCATCCCCTGGGCTCCC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGTGCCACCCCGAGGAGCTGGTCGGAGCTCTCTCGGAACACTCTCTGGGCATCCCCTGGGCTCCC 180

Qy 181 CTGAGCAGCTGCCCAGCCAGCCCTGCAGGTGGCAGGCCCTGGAAGGATCCCCGAGTTGGGT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CTGAGCAGCTGCCCAGCCAGCCCTGCAGGTGGCAGGCCCTGGAAGGATCCCCGAGTTGGGT 240

Qy 241 GGCCTTTCCTCTACCAGGGGCTCGCCAGCCCTTGCCACCACCATCTCGGCCAGCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GGCCTTTCCTCTACCAGGGGCTCGCCAGCCCTTGCCACCACCATCTCGGCCAGCAG 300

Qy 301 CCCACCTTGGACACACTGCAGCTGACGTGCGCCGACTTTGCCACCACATCTCGGCCAGCAG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CCCACCTTGGACACACTGCAGCTGACGTGCGCCGACTTTGCCACCACATCTCGGCCAGCAG 360

Qy 361 ATGGAAGAACTGGAATGGCCCTGCCCTGCGAGCCACCCAGGGTGCCATGCGGCCCTTC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 ATGGAAGAACTGGAATGGCCCTGCCCTGCGAGCCACCCAGGGTGCCATGCGGCCCTTC 420

Qy 421 GCCTCTGCTTCCAGCGCCGGGCAGGAGGGTCCTAGTTGCCTCCCCATCTGCAGAGCTTC 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GCCTCTGCTTCCAGCGCCGGGCAGGAGGGTCCTCCTAGTTGCCTCCCCATCTGCAGAGCTTC 480

Qy 481 CTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT 523
        ||||||||||||||||||||||||||||||||||||||||||
Db 481 CTGGAGGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT 523

```
Db  301 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        CCCACCTTGGACACACTGGAATGGCCCCTGACGTCGCCGACTTGCCACCACCATCTGGCAGCAG 360
Qy  361 ATGGAAGAACTGGGAATGGCCCCTGCCCCTGCAGCGCCAGGAGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ATGGAAGAACTGGGAATGGCCCCTGCCCCTGCAGCGCCAGGAGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 420
Qy  421 GCCTCTGCTTCCAGCGCCGCGGCCGGCAGGAGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCCTCTGCTTCCAGCGCCGCGGCCGGCAGGAGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
Qy  481 CTGGAGGTGTCGTACCGCGGTTCTACGCCGCCACCTTGCCCAGCCCT 523
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CTGGAGGTGTCGTACCGCGGTTCTACGCCGCCACCTTGCCCAGCCCT 523
```

RESULT 3
US-08-783-288-1
; Sequence 1, Application US/08783288
; Patent No. 5795968
; GENERAL INFORMATION:
;   APPLICANT: KUGA, TETSURO
;   APPLICANT: MIYAJI, HIROMASA
;   APPLICANT: SATO, MORIYUKI
;   APPLICANT: OKABE, MASAMI
;   APPLICANT: MORIMOTO, MAKOTO
;   APPLICANT: ITOH, SEIGA
;   APPLICANT: YAMASAKI, MOTOO
;   APPLICANT: YOKOO, YOSHIHARU
;   APPLICANT: YAMAGUCHI, KAZUO
;   APPLICANT: YOSHIDA, HAJIME
;   APPLICANT: YOSHINORI, KOMATSU
;   TITLE OF INVENTION: NOVEL POLYPEPTIDES
;   NUMBER OF SEQUENCES: 61
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: NIXON & VANDERHYE P.C.
;   STREET: 1100 NORTH GLEBE ROAD
;   CITY: ARLINGTON
;   STATE: VIRGINIA
;   COUNTRY: U.S.A.
;   ZIP: 22201-4714
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/783,288
;   FILING DATE: 10-JAN-1997
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/434,411
;   FILING DATE: 03-MAY-1995
;   APPLICATION NUMBER: JP 306799/86
;   FILING DATE: 23-DEC-1986
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 51357/88
;   FILING DATE: 04-MAR-1988
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 80088/88
;   FILING DATE: 31-MAR-1988
;   ATTORNEY/AGENT INFORMATION:
;   NAME: CRAWFORD, ARTHUR
;   REGISTRATION NUMBER: 25327
;   REFERENCE/DOCKET NUMBER: 249-73
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703)816-4000
;   TELEFAX: (703)816-4100
;   TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 525 base pairs
;   TYPE: nucleic acid

;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..522
;   OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
        STIMULATING FACTOR"
;   OTHER INFORMATION:
US-08-783-288-1

Query Match   100.0%; Score 523; DB 1; Length 525;
Best Local Similarity   100.0%; Pred. No. 5.5e-115;
Matches 523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
Qy   61 GTGAGGAAGATCCAGGGCCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
Qy  121 CTGTGCCACCCCGGAGGAGTCGGTGGTGCTCGGACACTCTCGGCATCCCTCGGGCTCCC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGTGCCACCCCGGAGGAGTCGGTGGTGCTCGGACACTCTCGGCATCCCTCGGGCTCCC 180
Qy  181 CTGAGCAGTGCCCCAGCCCTGCTTGAGCCAGGTGCAGGCTGCTGCCAACTCCATAGC 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGAGCAGTGCCCCAGCCCTGCTTGAGCCAGGTGCAGGCTGCTGCCAACTCCATAGC 240
Qy  241 GGCCTTTCCTTACCAGGGGATCCTCGCAGGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGCCTTTCCTTACCAGGGGATCCTCGCAGGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
Qy  301 CCCACCTTGGACACACTGGAGTCGAGCTGGACGTCGCCGACTTGCCACCACCATCTGGCAGCAG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CCCACCTTGGACACACTGGAGTCGAGCTGGACGTCGCCGACTTGCCACCACCATCTGGCAGCAG 360
Qy  361 ATGGAAGAACTGGGAATGGCCCCTGCCCCAGGGTGCATGCGCCTTC 420
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ATGGAAGAACTGGGAATGGCCCCTGCCCCAGGGTGCATGCGCCTTC 420
Qy  421 GCCTCTGCTTCCAGCGCCGGGGCGGGGTCCTCCAGTTGCTTCCCATCTGCAGAGCTTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCCTCTGCTTCCAGCGCCGGGGCGGGGTCCTCCAGTTGCTTCCCATCTGCAGAGCTTC 480
Qy  481 CTGGAGGTGTCGTACCGCGGTTCTACGCCGCCACCTTGCCCAGCCCT 523
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CTGGAGGTGTCGTACCGCGGTTCTACGCCGCCACCTTGCCCAGCCCT 523
```

RESULT 4
US-08-890-640-1
; Sequence 1, Application US/08890640
; Patent No. 5994518
; GENERAL INFORMATION:
;   APPLICANT: KUGA, TETSURO
;   APPLICANT: MIYAJI, HIROMASA
;   APPLICANT: SATO, MORIYUKI
;   APPLICANT: OKABE, MASAMI
;   APPLICANT: MORIMOTO, MAKOTO
;   APPLICANT: ITOH, SEIGA
;   APPLICANT: YAMASAKI, MOTOO
;   APPLICANT: YOKOO, YOSHIHARU
;   APPLICANT: YAMAGUCHI, KAZUO
;   APPLICANT: YOSHIDA, HAJIME
;   APPLICANT: YOSHINORI, KOMATSU
;   TITLE OF INVENTION: NOVEL POLYPEPTIDES
;   NUMBER OF SEQUENCES: 61
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: NIXON & VANDERHYE P.C.
;   STREET: 1100 NORTH GLEBE ROAD
;   CITY: ARLINGTON

```
Qy  361 ATGGAGGAACTGGGAATGCCCGGCGCGGCCCCTGCAGCCCACCAGGGTGCCATGCCGGCCTTC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ATGGAGGAACTGGGAATGCCCGGCGCGGCCCCTGCAGCCCACCAGGGTGCCATGCCGGCCTTC 420

Qy  421 GGCCTCTGCTTTCCAGCGCGCCGGGCGAGGAGGGTCCTCAGTTGCTCTCCCATCTGCAGAGCTTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGCCTCTGCTTTCCAGCGCGCCGGGCGAGGAGGGTCCTCAGTTGCTCTCCCATCTGCAGAGCTTC 480

Qy  481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
            ||||||||||||||||||||||||||||||||||||||||||||
Db  481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
```

RESULT 5
5194592-25
;Patent No. 5194592
; APPLICANT: YOSHIDA, HAJIME
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO NOVEL
;POLYPEPTIDES DERIVTIVES OF HUMAN GRANULOCYTE COLONY
;STIMULATING FACTOR
; NUMBER OF SEQUENCES: 83
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/318,527
; FILING DATE: 3-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 136,647
; FILING DATE: 22-DEC-1987
;SEQ ID NO:25:
; LENGTH: 525
5194592-25

Query Match   100.0%; Score 523; DB 6; Length 525;
Best Local Similarity   100.0%; Pred. No. 5.5e-115;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCAGCTCCTGCTCAAGTGCTTAGAGCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCAGCTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGGCCATGGGCCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGGCCATGGGCCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy  121 CTGTGCCACCCCGAGGAGGTGGTGCTGCTGCAGGCACTCTCTGGGCATCCCCTGGGCTCCC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGTGCCACCCCGAGGAGGTGGTGCTGCTGCAGGCACTCTCTGGGCATCCCCTGGGCTCCC 180

Qy  181 CTGAGCAGCTGCCCCAGCCCTGGCAGGTGCTGCTTGAGCCAACTCCATAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGAGCAGCTGCCCCAGCCCTGGCAGGTGCTGCTTGAGCCAACTCCATAGC 240

Qy  241 GGCCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGATCTCCCCGAGTTGGGT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGCCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGATCTCCCCGAGTTGGGT 300

Qy  301 CCCACCTTGGACACACTGCAGTCTGGACAGTGCCAGCGTGCCGACTTGCCACCATCTGCAGCAG 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CCCACCTTGGACACACTGCAGTCTGGACAGTGCCAGCGTGCCGACTTGCCACCATCTGCAGCAG 360

Qy  361 ATGGAGAGAACTGGGAATGCCCGGCGCGGCCCCTGCAGCCCACCAGGGTGCCATGCCGGCCTTC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ATGGAGAGAACTGGGAATGCCCGGCGCGGCCCCTGCAGCCCACCAGGGTGCCATGCCGGCCTTC 420

Qy  421 GGCCTCTGCTTTCCAGCGCGCCGGGCGAGGAGGGTCCTCAGTTGCTCTCCCATCTGCAGAGCTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGCCTCTGCTTTCCAGCGCGCCGGGCGAGGAGGGTCCTCAGTTGCTCTCCCATCTGCAGAGCTTC 480

Qy  481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
           ||||||||||||||||||||||||||||||||||||||||||||
Db  481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
```

RESULT 6

STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/890,640
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/434,411
FILING DATE: 03-MAY-1995
APPLICATION NUMBER: JP 306799/86
FILING DATE: 23-DEC-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 51357/88
FILING DATE: 04-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 80088/88
FILING DATE: 31-MAR-1988
ATTORNEY/AGENT INFORMATION:
NAME: CRAWFORD, ARTHUR
REGISTRATION NUMBER: 25327
REFERENCE/DOCKET NUMBER: 249-73
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100
TELEX: 200797 NIXN UR
US-08-890-640-1
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 525 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..522
OTHER INFORMATION: /product= "HUMAN GRANULOCYTE COLONY
OTHER INFORMATION: STIMULATING FACTOR"

Query Match   100.0%; Score 523; DB 2; Length 525;
Best Local Similarity   100.0%; Pred. No. 5.5e-115;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCAGCTCCTGCTCAAGTGCTTAGAGCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCAGCTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGGCCATGGGCCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGGCCATGGGCCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy  121 CTGTGCCACCCCGAGGAGGTGGTGCTGCTGCGGGCCCTGGCATCCCCTGGGCTCCC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGTGCCACCCCGAGGAGGTGGTGCTGCTGCGGGCCCTGGCATCCCCTGGGCTCCC 180

Qy  181 CTGAGCAGCTGCCCCAGCCCTGGCAGCTGGCGCTGCTTGAGCCAACTCCATAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGAGCAGCTGCCCCAGCCCTGGCAGCTGGCGCTGCTTGAGCCAACTCCATAGC 240

Qy  241 GGCCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGCCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300

Qy  301 CCCACCTTGGACACACTGCAGTCTGGACAGTGCCAGCGTGCCGACTTGCCACCATCTGCAGCAG 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CCCACCTTGGACACACTGCAGTCTGGACAGTGCCAGCGTGCCGACTTGCCACCATCTGCAGCAG 360
```

```
US-08-797-689-13
; Sequence 13, Application US/0879769
; Patent No. 5876969
  GENERAL INFORMATION:
     APPLICANT: Fleer, Reinhard
     APPLICANT: Fournier, Alain
     APPLICANT: Guitton, Jean-Dominique
     APPLICANT: Jung, Gerard
     APPLICANT: Yeh, Patrice
     TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
     TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
     TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
     NUMBER OF SEQUENCES: 36
     CORRESPONDENCE ADDRESS:
        ADDRESSE: Rhone-Poulenc Rorer Inc.
        STREET: 500 Arcola Road, 3C43
        CITY: Collegeville
        STATE: PA
        COUNTRY: USA
        ZIP: 19426
  COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: Macintosh
     OPERATING SYSTEM: System 7.1
     SOFTWARE: Word 5.1 (PatentIn)
  CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/797,689
     FILING DATE: 31-JAN-1997
     CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
     APPLICATION NUMBER: US 08/256,927
     FILING DATE: 28-JUL-1994
     APPLICATION NUMBER: FR 92/01064
     FILING DATE: 31-JAN-1992
  PRIOR APPLICATION DATA:
     APPLICATION NUMBER: PCT/FR93/00085
     FILING DATE: 28-JAN-1993
  ATTORNEY/AGENT INFORMATION:
     NAME: Smith Ph.D., Julie K.
     REGISTRATION NUMBER: P-38,619
     REFERENCE/DOCKET NUMBER: ST92006-US
  TELECOMMUNICATION INFORMATION:
     TELEPHONE: (610) 454-3839
     TELEFAX: (610) 454-3808
  INFORMATION FOR SEQ ID NO: 13:
     SEQUENCE CHARACTERISTICS:
        LENGTH: 541 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: double
        TOPOLOGY: linear
     MOLECULE TYPE: cDNA
     FEATURE:
        NAME/KEY: CDS
        LOCATION: 3..536
US-08-797-689-13

Query Match         97.6%;  Score 510.2;  DB 2;  Length 541;
Best Local Similarity 98.5%;  Pred. No. 5.9e-112;
Matches 515;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy         1 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  71

Qy        61 GTGAGGAAGATCCAGGGCGATGGCGAGGAAGAGCTGTGTGCCACCTACAAG  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        72 GTGAGGAAGATCCAGGGCGATGGCGAGGAAGAGCTGTGTGCCACCTACAAG  131

Qy       121 CTGTGCCACCCCGAGGAGTGGTGTCTGCTTCGGACACTCTCTGGGCATCTCTGGGCTCCC  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       132 CTGTGCCACCCCGAGGAGTGGTGTCTGCTTCGGACACTCTCTGGGCATCTCTGGGCTCCC  191

Qy       181 CTGAGCAGCAGCTGCGCCCCAGCCAGGCCCTGCAGCTGCGCAGGCTGCTTGAGCCAACTCCATAGC  240
```

```
Db       192 CTGAGCTTCTGCCCCAGCCAGGCCCTGCGCAGCAGCGGCTGCTGCAGCAGCAGGATCTCCCCGAGTTGGGT  251
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       241 GGCCTTTTCCTCTACCAGGGCTCCTGCAGGGGC CTGGAAGGATCCCCCGAGTTGGGT  300
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       252 GGCCTTTTCCTCTACCAGGGGTCCTGCAGGGCCCTGGAAGGATATCCCCGAGTTGGGT  311
Qy       301 CCCACCCTGGACACACTGCAGCTGCGCCGACTTTGCCACCACATCTGCAGCAG  360
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       312 CCCACCCTGGACACACTGCAGCTGGACGTCTGCCGACTTTGCCACCACATCTGCAGCAG  371
Qy       361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCGCCACCAGGGTGCCAGGCCTTC  420
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       372 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCGCCACCAGGGTGCCATGCGGCCTTC  431
Qy       421 GCCTCCTCTTTCCAGCGCGCCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC  480
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       432 GCCTCGTCTTTCCAGCGCGCCGGGCAGGAGGGGTCCTCGTGTTGCTAGCCATCTGCAGAGCTTC  491
Qy       481 CTGGAGGTGTCGTACCGGCGTTCTACCGCCACCTTGCCCAGCCCT  523
              ||||||||||||||||||||||||||||||||||||||
Db       492 CTGGAGGTGTCGTACCGGCGTTCACGCCACCTTGCCGCAGCCCT  534
```

```
RESULT 7
US-09-984-186-13
; Sequence 13, Application US/09984186
; Patent No. 6686179
  GENERAL INFORMATION:
     APPLICANT: Fleer, Reinhard
                Fournier, Alain
                Guitton, Jean-Dominique
                Jung, Gerard
                Yeh, Patrice
     TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
                         PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
                         CONTAINING SAID POLYPEPTIDES
  NUMBER OF SEQUENCES: 36
  CORRESPONDENCE ADDRESS:
     ADDRESSEE: Rhone-Poulenc Rorer Inc.
     STREET: 500 Arcola Road, 3C43
     CITY: Collegeville
     STATE: PA
     COUNTRY: USA
     ZIP: 19426
  COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: Macintosh
     OPERATING SYSTEM: System 7.1
     SOFTWARE: Word 5.1 (PatentIn)
  CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/09/984,186
     FILING DATE: 29-Oct-2001
     CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
     APPLICATION NUMBER: US/08/797,689
     FILING DATE: 31-JAN-1997
     APPLICATION NUMBER: US 08/256,927
     FILING DATE: 28-JUL-1994
     APPLICATION NUMBER: FR 92/01064
     FILING DATE: 31-JAN-1992
     APPLICATION NUMBER: PCT/FR93/00085
     FILING DATE: 28-JAN-1993
  ATTORNEY/AGENT INFORMATION:
     NAME: Smith Ph.D., Julie K.
     REGISTRATION NUMBER: P-38,619
     REFERENCE/DOCKET NUMBER: ST92006-US
  TELECOMMUNICATION INFORMATION:
     TELEPHONE: (610) 454-3839
     TELEFAX: (610) 454-3808
  INFORMATION FOR SEQ ID NO: 13:
     SEQUENCE CHARACTERISTICS:
        LENGTH: 541 base pairs
```

```
          TYPE: nucleic acid
          STRANDEDNESS: double
          TOPOLOGY: linear
       MOLECULE TYPE: cDNA
       FEATURE:
          NAME/KEY: CDS
          LOCATION: 3..536
       SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-984-186-13

Query Match     97.6%;     Score 510.2;   DB 4;  Length 541;
Best Local Similarity  98.5%;    Pred No. 5.9e-112;
Matches  515;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy    1 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 71

Qy   61 GTGAGGAAGATCCAGGGCGCAGGCGCCATGGGCGCAGAGCTGTGTGCCACCTACAAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72 GTGAGGAAGATCCAGGGCGCAGGCGCCATGGGCGCAGAGCTGTGTGCCACCTACAAG 131

Qy  121 CTGTGCCACCCCCGAGGAGCTGGTGCTGCTCGGCGACACTCTCTGGGCTCCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  132 CTGTGCCACCCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCTCCC 191

Qy  181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCTTGAGCCAACTCCATAGC 240
          ||||||||||||||||||||||||||||||||||||||||||
Db  192 CTGAGCTCTGCCCCAGCCAGGCCCTGCTTGAGCCAACTCCATAGC 251

Qy  241 GGCCTTTCCTCTACCAGGGGATCTCCGGAAGGGATCTCCCCGAGTTGGGT 300
          |||||||||||||||||||||||||||||||||||||||||
Db  252 GGCCTTTTCCTCTACCAGGGGTCTCGCAGCCCTGGAAGGGATATCCCCGAGTTGGGT 311

Qy  301 CCCACCTTGGACACACTGAGCTGCGACGTCGCGACTTTGCCACCATCTGGCAGCAG 360
          ||||||||||||||||||||||||||||||||||||||||
Db  312 CCCACCTTGGACACACTGAGCTGCGACGTCGCGACTTTGCCACCATCTGGCAGCAG 371

Qy  361 ATGGAAGAACTGGGAATGGCCCCTGCCCCTGCAGCCCACCAGGGTGCGCGGCCTTC 420
          ||||||||||||||||||||||||||||||||||||||||
Db  372 ATGGAAGAACTGGGAATGGCCCCTGCCCCTGCAGCCCACCAGGGTGCCATGCCGCCTTC 431

Qy  421 GCCTCTGCTTTCCAGCGCGCCGGCGAGGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
          ||||||||||||||||||||||||||||||||||||||||
Db  432 GCCTCTGCTTTCCAGCGCGCCGGCGAGGGGGTCCTGGTTGCTAGCCATCTGCAGAGCTTC 491

Qy  481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
          |||||||||||||||||||||||
Db  492 CTGGAGGTGTCGTACCGCGGTTCTACGCCACCTTGCCGCAGCCCT 534
```

RESULT 8
US-08-256-938-1
; Sequence 1, Application US/08256938
; Patent No. 5665863
GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Rhone-Poulenc Rorer Inc.
;   STREET: 500 Arcola Road, 3C43
;   CITY: Collegeville
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19426
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: Macintosh
;   OPERATING SYSTEM: System 7.1
;   SOFTWARE: Word 5.0 (PatentIn)

```
CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/08/256,938
   FILING DATE:
   CLASSIFICATION: 435
PRIOR APPLICATION DATA:
   APPLICATION NUMBER: FR 92/01065
   FILING DATE: 31-JAN-1992
ATTORNEY/AGENT INFORMATION:
   NAME: Goodman, Rosanne
   REGISTRATION NUMBER: 32,534
   REFERENCE/DOCKET NUMBER: ST92007-US
TELECOMMUNICATION INFORMATION:
   TELEPHONE: (610) 454-3817
   TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 1:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 2382 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
   MOLECULE TYPE: cDNA
   HYPOTHETICAL: NO
   ANTI-SENSE: NO
   FEATURE:
      NAME/KEY: CDS
      LOCATION: 26..2377
   FEATURE:
      NAME/KEY: misc_feature
      LOCATION: 1842..1848
      OTHER INFORMATION:  /label= MstII-site
   FEATURE:
      NAME/KEY: misc_feature
      LOCATION: 1861..1866
      OTHER INFORMATION:  /label= ApaI-site
   FEATURE:
      NAME/KEY: misc_feature
      LOCATION: 2035..2040
      OTHER INFORMATION:  /label= SstI-site
US-08-256-938-1

Query Match       97.6%;      Score 510.2;   DB 1;  Length 2382;
Best Local Similarity  98.5%;      Pred. No. 8.2e-112;
Matches  515;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy     1 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1853 ACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 1912

Qy    61 GTGAGGAAGATCCAGGGCCATGGGCGCAGAGCTGTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1913 GTGAGGAAGATCCAGGGCCAGGCCGCCATGGGCGCAGAGCTGTGTGCCACCTACAAG 1972

Qy   121 CTGTGCCACCCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCCATCCCCTGGGCTCCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1973 CTGTGCCACCCCCGAGGAGCTGGTGCTGCTCGGCGACACTCTCTGGGCCATCCCCTGGGCTCCC 2032

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCTTGAGCCAACTCCATAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||
Db  2033 CTGAGCTCCTGCCCCAGCCAGGCCCTGCTTGAGCCAACTCCATAGC 2092

Qy   241 GGCCTTTCCTCTACCAGGGGATCTCCGGAAGGGATCTCCCCGAGTTGGGT 300
           ||||||||||||||||||||||||||||||||||||||||||||
Db  2093 GGCCTTTTCCTCTACCAGGGGTCTCGCAGCCCTGGAAGGGATATCCCCGAGTTGGGT 2152

Qy   301 CCCACCTTGGACACACTGAGCTGCGACGTCGCGACTTTGCCACCATCTGGCAGCAG 360
           |||||||||||||||||||||||||||||||||||||||||||
Db  2153 CCCACCTTGGACACACTGAGCTGCGACGTCGCGACTTTGCCACCACCATCTGGCAGCAG 2212

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCCTGCAGCCCACCAGGGTGCCATGCCGGCCTTC 420
           ||||||||||||||||||||||||||||||||||||||||||||
Db  2213 ATGGAAGAACTGGGAATGGCCCCTGCCCCTGCAGCCCACCAGGGTGCCATGCCGGCCTTC 2272

Qy   421 GCCTCTGCTTTCCAGCGCCGGCGAGGAGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480
```

```
Db   2273  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           GCCTCTGCTTCCAGCCGCCGGGCAGGAGGGTCCTGGTTGCTAGCCATCTGCAGAGCTTC 2332

Qy    481  CTGGAGGTGTCGTACCCGTTCTACGCGCACCTTGCCCAGCCCT 523
           |||||||||||||||||||||||||| ||||| |||||||||||
Db   2333  CTGGAGGTGTCGTACCCGTTCTACGCGCACCTTGCGCAGCCCT 2375
```

RESULT 9
US-08-256-938-3
; Sequence 3, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
;   APPLICANT: Yeh, Patrice
;   TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
;   TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
;   TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
;   NUMBER OF SEQUENCES: 12
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Rhone-Poulenc Rorer Inc.
;   STREET: 500 Arcola Road, 3C43
;   CITY: Collegeville
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19426
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: Macintosh
;   OPERATING SYSTEM: System 7.1
;   SOFTWARE: Word 5.0 (PatentIn)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/256,938
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: FR 92/01065
;   FILING DATE: 31-JAN-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Goodman, Rosanne
;   REGISTRATION NUMBER: 32,534
;   REFERENCE/DOCKET NUMBER: ST92007-US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (610) 454-3817
;   TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2455 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 26..2389
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 106..111
;   OTHER INFORMATION: /label= ApaI-site
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 280..285
;   OTHER INFORMATION: /label= SstI-site
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 2376..2382
;   OTHER INFORMATION: /label= MstII-site
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 26..97
;   FEATURE:
;   NAME/KEY: misc_feature

;   LOCATION: 620..631
;   OTHER INFORMATION: /label= polyGly-linker
; US-08-256-938-3

   Query Match   97.4%; Score 509.2; DB 1; Length 2455;
   Best Local Similarity   98.5%; Pred. No. 1.4e-111;
   Matches 514; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

```
Qy     1  ACCCCCTGGGCCTCCTGCCAGCTCCCTGCCCAGAGCTTCTGCTCAAGTGCTTAGAGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    98  ACCCCCTGGGCCTCCTGCCAGCTCCCTGCCCAGAGCTTCTGCTCAAGTGCTTAGAGCAA 157

Qy    61  GTGAGGAAGATCCAGGGCGGCGATGGCGCCAGCCCTCCAGGAGAGAGCGTGTGCCACCTACAAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   158  GTGAGGAAGATCCAGGGCGGCGATGGCGCCAGCCCTCCAGGAGAGAGCGTGTGCCACCTACAAG 217

Qy   121  CTGTGCCACCCCGGAGGAGCTGGTGCTGCTGGACACTCTCTGGCATCCCTGGCTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   218  CTGTGCCACCCCGGAGGAGCTGGTGCTGCTGGACACTCTCTGGCATCCCTGGCTCCC 277

Qy   181  CTGAGCAGCTGCCCAGCCAGGCCCTGGCCAGCTGGCAGGCTGCTTGGCCAACTCCATAGC 240
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   278  CTGAGCTCTGCCCAGCCAGGCCCTGGCCAGCTGGCAGGCTGCTTGGCCAACTCCCATAGC 337

Qy   241  GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCGAGTTGGGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   338  GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATATCCCCGAGTTGGGT 397

Qy   301  CCCACCTGGACACACTGCAGCTGCGACGTGCGCCACCATCTGCGCAGCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   398  CCCACCTTGGACACACTGCAGCTGCGACGTGCGCCACCATCTGCGCAGCAG 457

Qy   361  ATGGAAGAACTGGGAATGGCCCCTGCCTCCCAGGGTGCCACCCAGGGTGCCGGCCTTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   458  ATGGAAGAACTGGGAATGGCCCCTGCCTCCCAGGGTGCCACCCAGGGTGCCGGCCTTC 517

Qy   421  GCCTCTGCTTCCAGCCGCCGGGCAGGAGGGTCCTAGTTGCTCCCATCTGCAGAGCTTC 480
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   518  GCCTCTGCTTCCAGCCGCCGGGCAGGAGGGTCCTGGTTGCTAGCCATCTGCAGAGCTTC 577

Qy   481  CTGGAGGTGTCGTACCCGTTCTACGCGCACCTTGCCCAGCCC 522
          |||||||||||||||||||||||||||||||||||||||||||
Db   578  CTGGAGGTGTCGTACCCGTTCTACGCGCACCTTGCGCAGCCC 619
```

RESULT 10
US-08-797-689-15
; Sequence 15, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
;   APPLICANT: Fleer, Reinhard
;   APPLICANT: Fournier, Alain
;   APPLICANT: Guitton, Jean-Dominique
;   APPLICANT: Jung, Gerard
;   APPLICANT: Yeh, Patrice
;   TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;   TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;   TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
;   NUMBER OF SEQUENCES: 36
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Rhone-Poulenc Rorer Inc.
;   STREET: 500 Arcola Road, 3C43
;   CITY: Collegeville
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19426
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: Macintosh
;   OPERATING SYSTEM: System 7.1
;   SOFTWARE: Word 5.1 (PatentIn)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/797,689

; Patent No. 6686179
; GENERAL INFORMATION:
;   APPLICANT: Fleer, Reinhard
;   Fournier, Alain
;   Guitton, Jean-Dominique
;   Jung, Gerard
;   Yeh, Patrice
;   TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;   PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;   CONTAINING SAID POLYPEPTIDES
;   NUMBER OF SEQUENCES: 36
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Rhone-Poulenc Rorer Inc.
;   STREET: 500 Arcola Road, 3C43
;   CITY: Collegeville
;   STATE: PA
;   COUNTRY: USA
;   ZIP: 19426
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: Macintosh
;   OPERATING SYSTEM: System 7.1
;   SOFTWARE: Word 5.1 (PatentIn)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/984,186
;   FILING DATE: 29-Oct-2001
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/797,689
;   FILING DATE: 31-JAN-1997
;   APPLICATION NUMBER: US 08/256,927
;   FILING DATE: 28-JUL-1994
;   APPLICATION NUMBER: FR 92/01064
;   FILING DATE: 31-JAN-1992
;   APPLICATION NUMBER: PCT/FR93/00085
;   FILING DATE: 28-JAN-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Smith Ph.D., Julie K.
;   REGISTRATION NUMBER: P-38,619
;   REFERENCE/DOCKET NUMBER: ST92006-US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (610) 454-3839
;   TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 15:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2455 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 26..2389
;   SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-984-186-15

Query Match   97.4%; Score 509.2; DB 4; Length 2455;
Best Local Similarity 98.5%; Pred. No. 1.4e-111;
Matches 514; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy   1 ACCCCCTGGGCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   98 ACCCCCTGGGCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 157

Qy   61 GTGAGGAAGATCCAGGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   158 GTGAGGAAGATCCAGGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 217

Qy   121 CTGTGCCACCCCGAGGAGTGGTGCTGCTCGGACACTCTCTGGGCTCCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   218 CTGTGCCACCCCGAGGAGTGGTGCTGCTCGGACACTCTCTGGGCTCCC 277

Qy   181 CTGAGCAGCTGCCCCAGCCAGCCCTGCAGGTGCCAGCGTGCTTGAGCAACTCCATAGC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   278 CTGAGCTGCTGCCCCAGCCAGCCCTGCAGGTGCCAGCGTGCTTGAGCAACTCCATAGC 337

US-09-984-186-15
; Sequence 15, Application US/09984186

```
                                                  INDIVIDUAL ISOLATE:
                                                  DEVELOPMENTAL STAGE:
                                                  HAPLOTYPE:
                                                  TISSUE TYPE:
                                                  CELL TYPE:
                                                  CELL LINE:
                                                  ORGANELLE:
                                              IMMEDIATE SOURCE:
                                                  LIBRARY:  Fred Hutchinson Cancer Research Center
                                                  CLONE:
                                              POSITION IN GENOME:
                                                  CHROMOSOME/SEGMENT:
                                                  MAP POSITION:
                                                  UNITS:
                                              FEATURE:
                                                  NAME/KEY:  G-CSF
                                                  LOCATION:
                                                  IDENTIFICATION METHOD:
                                                  OTHER INFORMATION:
                                              PUBLICATION INFORMATION:
                                                  AUTHORS:
                                                  TITLE:
                                                  JOURNAL:
                                                  VOLUME:
                                                  ISSUE:
                                                  PAGES:
                                                  DATE:
                                                  DOCUMENT NUMBER:
                                                  FILING DATE:
                                                  PUBLICATION DATE:
                                                  RELEVANT RESIDUES IN SEQ ID NO:  1:
                                              US-08-149-101A-1

Query Match                 96.9%;   Score 507;  DB 3;  Length 525;
Best Local Similarity   99.1%;   Pred. No. 3.3e-111;
Matches  513;   Conservative  0;  Mismatches  10;  Indels  0;  Gaps  0;

Qy      1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAGTGCTTAGAGCAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAGTGTGTGAGCAA 60

Qy     61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTTCAGGAGAGAAGCGTGTGCCACTACAAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCTTCAGGAGAGAAGCGTGTTAGCCACTACAAG 120

Qy    121 CTGTGCCACCCCGAGGAGTCGGTGCTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGTGCCACCCCGAGGAGTCGGTGCTGCTCGGACACTCTCTGGGCATCCCTGGGCTCCC 180

Qy    181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCCAGGCTGCTTGAGCCAACTCCATAGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCCAGGCTGCTTGAGCCAACTCCATAGC 240

Qy    241 GGCCTTTTCCTCTACCAGGGGCTCCTGCCAGGCCCTGAAGGATCTCCCCGAGTTGGGT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCCTTTTCCTCTACCAGGGGCTCCTGCCAGGCCCTGAAGGATCTCCCCGAGTTGGGT 300

Qy    301 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTTGCCACCACCATCTGCAGCAAG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTTGCCACCACCAGGTGCAGCAG 360

Qy    361 ATGGAAGAACTGGGAATGGCCCTGCCTGCGCTCGAGCCCCAGGGTGCCATGCCGGCCTTC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATGGAAGAACTGGGAATGGCCCTGCCTGCGCTCGAGCCCCAGGGTGCCATGCCGGCCTTC 420

Qy    421 GCTCTGCTTTCCAGCGCGCCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCTCTGCTTTCCAGCGCGCCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480

Qy    481 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
          |||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
```

```
Db    278 CTGAGCTCCTGCCCCAGCCAGGCCCTGCGAGCTGCCAGGCCCTGCTTGAGCCAACTCCATAGC 337
          |||||||||||   |||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 GGCCTTTTCCTCTACCAGGGGCTCCTGCCAGGCCCTGAAGGATCTCCCCGAGTTGGGT 300
          |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db    338 GGCCTTTTCCTCTACCAGGGGCTCCTGCCAGGCCCTGCGAGGGATATCCCCGAGTTGGGT 397
Qy    301 CCCACCTTGGACACACTGACGTGCAGCTGCCGACCTTTGCCACCACCATCTGGCAGCAG 360
          |||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db    398 CCCACCTTGGACACACTGCAGCTGGACGTGCCGACCTTTGCCACCACCATCTGGCAGCAG 457
Qy    361 ATGGAAGAACTGGGAATGGCCCTGCCCCAGCCCACCAGGGTGCCATGCCGGCCTTC 420
          |||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db    458 ATGGAAGAACTGGGAATGGCCCTGCGAGCCCACCAGGGTGCCATGCCGGCCTTC 517
Qy    421 GCTCTGCTTTCCAGCGCGCCGGGGGTCCTAGTTGCTCCCATCTGCAGAGCTTC 480
          ||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db    518 GCTCTGCTTTCCAGCGCGCCGGGGGAGGGTGCTAGCCATCTGCTAGCAGAGCTTC 577
Qy    481 CTGGAGGTGTCGTACCGCGTTCTACGCGCCAGCCC 522
          ||||||||||||||||||||||||   ||||||||
Db    578 CTGGAGGTGTCGTACCGCGTTCTACGCCCCACCTTGCGCAGCCC 619
```

```
RESULT 12
US-08-149-101A-1
; Sequence 1, Application US/08149101A
; Patent No. 6171824
; GENERAL INFORMATION:
    APPLICANT:  Todaro, George J.
    APPLICANT:  Leung, David W.
    APPLICANT:  Rose, Timothy M.
    TITLE OF INVENTION:  HYBRID CYTOKINES
    NUMBER OF SEQUENCES:  26
    CORRESPONDENCE ADDRESS:
      ADDRESSE:  Cell Therapeutics, Inc.
      STREET:  200 Elliott Avenue West, Suite 400
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  U.S.A.
      ZIP:  98119
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  3.5" disk, 1.44Mb, double side, high density
      COMPUTER:  AST-IBM Compatible
      OPERATING SYSTEM:  MS-DOS Version 6
      SOFTWARE:  WORD for WINDOWS
    CURRNT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/149,101A
      FILING DATE:  8-No. 6171824-1993
      CLASSIFICATION:  530
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US08/097,869
      FILING DATE:  27-Jul-1993
    ATTORNEY/AGENT INFORMATION:
      NAME:  Oster, Jeffrey B. and Faciszewski, Stephen
      REGISTRATION NUMBER:  32,585 and 36,131, respectively
      REFERENCE/DOCKET NUMBER:  0105A
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206)282-7100
      TELEFAX:  (206)284-6206
    INFORMATION FOR SEQ ID NO:  1:
      SEQUENCE CHARACTERISTICS:
        LENGTH:  525
        TYPE:  nucleic acid
        STRANDEDNESS:  double stranded
        TOPOLOGY:  linear
      MOLECULE TYPE:  cDNA to mRNA
      HYPOTHETICAL:  no
      ANTI-SENSE:  no
      FRAGMENT TYPE:
      ORIGINAL SOURCE:
        ORGANISM:  homo sapien
        STRAIN:
```

RESULT 13
PCT-US94-12873-1
; Sequence 1, Application PC/TUS9412873
; GENERAL INFORMATION:
;   APPLICANT:
;   APPLICANT:
;   APPLICANT:
;   TITLE OF INVENTION: HYBRID CYTOKINES
;   NUMBER OF SEQUENCES: 26
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5" diskette, 1.44Mb, double side, high density
;   COMPUTER: AST-IBM Compatible
;   OPERATING SYSTEM: MS-DOS Version 6
;   SOFTWARE: WORD for WINDOWS
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US94/12873
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 525
;   TYPE: nucleic acid
;   STRANDEDNESS: double stranded
;   TOPOLOGY: linear
PCT-US94-12873-1

Query Match   96.9%;   Score 507;   DB 5;   Length 525;
Best Local Similarity 98.1%;   Pred. No. 3.3e-111;
Matches 513;   Conservative   0;   Mismatches 10;   Indels 0;   Gaps 0;

Qy   1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTGTGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGCTTCCAGGAGAGCTGTGCCACCTACAAG 120
           ||||||||||||||||||| |||||||||||||||||||||||||| |||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGCTTCCAGGAGAGCTGTTAGCCACCTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGGACACTCTCTGGGCATCCCCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 240

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCTGCCAGCCCTGGAAGGGATCTCCCCCGAGTTGGGT 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTTCCTCTACCAGGGGCTCTGCCAGCCCTGGAAGGGATCTCCCCCGAGTTGGGT 300

Qy   301 CCCACCTTGGACACACTGGAGTCGCGACTTTGCCACCATCTGGCAGCAG 360
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACACTGGAGTCGCGACTTTGCCACCATCTGGCAGCAG 360

Qy   361 ATGGAAGAACTGGGAATGCCCCTGCCCTGCAGCCCAGGGTGCCATGCCGGCCTTC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGGAAGAACTGGGAATGCCCCTGCCCTGCAGCCCAGGGTGCCATGCCGGCCTTC 420

Qy   421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGTCCTCCCATCTGCGAGAGCTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGTCCTCCCATCTGCGAGAGCTTC 480

Qy   481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
           ||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523

RESULT 14
US-08-469-318-177
; Sequence 177, Application US/08469318
; Patent No. 6022535
; GENERAL INFORMATION:
;   APPLICANT:

;   TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
;   TITLE OF INVENTION: Protein
;   NUMBER OF SEQUENCES: 196
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/469,318
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/446,872
;   FILING DATE:
;   INFORMATION FOR SEQ ID NO: 177:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 546 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-469-318-177

Query Match   96.3%;   Score 503.8;   DB 3;   Length 546;
Best Local Similarity 97.7%;   Pred. No. 1.9e-110;
Matches 511;   Conservative   0;   Mismatches 12;   Indels 0;   Gaps 0;

Qy   1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 ACACCATTAGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 66

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGCTTCCAGGAGAGCTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 GTGAGGAAGATCCAGGGCGATGGGCAGCGCTTCCAGGAGAGCTGTGCCACCTACAAG 126

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGGACACTCTCTGGGCATCCCCTGGGCTCCC 186

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 CTGAGCTTCCTGCCCCAGCCAGGCCCTGCCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 246

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCTGCCAGCTGGACGTGACGACTTGGCCACCATCTGGCAGCAG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 GGCCTTTTCCTCTACCAGGGGCTCTGCCAGCTGGACGTGACGACTTGGCCACCATCTGGCAGCAG 306

Qy   301 CCCACCTTGGACACACTGGAGTCGCGACTTTGCCACCATCTGGCAGCAG 360
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 CCCACCTTGGACACACTGGAGTCGCGACTTTGCCACCATCTGGCAGCAG 366

Qy   361 ATGGAAGAACTGGGAATGCCCCTGCCCTGCAGCCCCAGGGTGCATGCCGGCCTTC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 ATGGAAGAACTGGGAATGCCCCTGCCCTGCAGCCCCAGGGTGCATGCCGGCCTTC 426

Qy   421 GCCTCTGCTTTCCAGCGCCGGGGGTCCTCCCATCTGCCATCTGGTTGCTAGCCATTCGAGAGCTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 GCCTCTGCTTTCCAGCGCCGGGGGTCCTCCCATCTGCCATCTGGTTGCTAGCCATTCGAGAGCTTC 486

Qy   481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
           ||||||||||||||||||||||||||||||||||||||||||||
Db   487 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 529

RESULT 15
US-08-468-609A-177
; Sequence 177, Application US/0846809A
; Patent No. 6030812
; GENERAL INFORMATION:
;   APPLICANT: Abrams, Mark A.
;   APPLICANT: Bauer, S. C.
;   APPLICANT: Braford-Goldberg, Sarah R.

; APPLICANT: Caparon, Maire H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olins, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Inteleukin-3 (IL-
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,609A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/192,325
; FILING DATE: 14-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2790/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6986
; TELEFAX: (314)737-6972
; INFORMATION FOR SEQ ID NO: 177:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-468-609A-177

Query Match   96.3%;   Score 503.8;   DB 3;   Length 546;
Best Local Similarity 97.7%;   Pred. No. 1.9e-110;
Matches 511; Conservative   0; Mismatches 12; Indels   0; Gaps   0;

Qy   1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
            ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   7 ACACCATTAGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 66

Qy   61 GTGAGGAAGATCACAGGGCGATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 GTGAGGAAGATCACAGGGCGATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 126

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 186

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 CTGAGCTCGCCCCAGCCAGGCCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 246

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGCCCTGGAAGGGATATCCCCGAGTTGGGT 306

Qy   301 CCCACCTTGGACACTGCGAGCTGCCGACTTGCCGACTTGCCACCATCTGGCAGCAG 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 CCCACCTTGGACACACTGCGAGCTGCCGACGTGCCGACTTGCCACCATCTGGCAGCAG 366

Qy   361 ATGGAGGAACTGGGAATGCCCGCCCTGCCAGCCACCAGGGTGCCATGCGCGGCCTTC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 ATGGAGGAACTGGGAATGCCCGCCCTGCCAGCCACCAGGGTGCCATGCGCGGCCTTC 426

Qy   421 GCCTCTGCTTCCAGCGCGCAGGAGGGGTCCTAGTTGCTCCCATCTGCAGAGCTTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 GCCTCTGCTTCCAGCGCGCGGGCGCAGGAGGGGTCCTGGTTGCTCCCATCTGCAGAGCTTC 486

Qy   481 CTGGAGGTGTCGTACCGGCGTTCACGGCCACCTTGCCCAGCCCT 523
            ||||||||||||||||||||||||||||||||||||||||||||
Db   487 CTGGAGGTGTCGTACCGGCGTTCTACGCCGGCTTCTACGCCACCTTGCCCAGCCCT 529

Search completed: October 9, 2004, 04:24:23
Job time : 83.3471 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2004, 21:19:03 ; Search time 397.272 Seconds
(without alignments)
5592.665 Million cell updates/sec

Title: US-10-009-792C-18_COPY_88_610
Perfect score: 523
Sequence: 1 acccccctgggccctgccag........acgccacctgcccagccct 523

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   N_Geneseq 29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	523	100.0	525	1	AAN80947	Aan80947 G-CSF gen
2	523	100.0	531	4	AAD19772	Aad19772 Human hG-
3	523	100.0	615	4	AAD19771	Aad19771 Human hG-
4	523	100.0	1520	1	AAN70223	Aan70223 Plasmid p
5	523	100.0	1520	1	AAN71320	Aan71320 Sequence
6	523	100.0	1521	1	AAN60937	Aan60937 Plasmid p
7	523	100.0	1525	1	AAN81478	Aan81478 Sequence
8	523	100.0	1525	1	AAN91086	Aan91086 Plasmid p
9	521.4	99.7	525	7	ACC78869	Acc78869 Human gra
10	521.4	99.7	549	3	ABA94177	Aba94177 Granulocy
11	521.4	99.7	1415	1	AAN71089	Aan71089 Sequence
12	521.4	99.7	1508	7	ABX63825	Abx63825 Human cDN
13	513.4	98.2	644	5	AAI71848	Aai71848 Recombina
14	510.2	97.6	540	2	AAZ21169	Aaz21169 Recombina
15	510.2	97.6	556	1	AAN90533	Aan90533 Synthetic
16	510.2	97.6	777	3	AAA71317	Aaa71317 Mammalian
17	510.2	97.6	780	2	AAA71312	Aaa71312 MF alpha1
18	510.2	97.6	2382	2	AAQ45987	Aaq45987 Prepro-HS
19	509.2	97.4	2455	2	AAQ45988	Aaq45988 G-CSF-(G1
20	508.6	97.3	1039	2	AAQ24283	Aaq24283 IL-3:G-CS
21	507.6	97.1	522	4	AAF56684	Aaf56684 Human gra
22	507	96.9	525	2	AAQ87151	Aaq87151 Human gra
23	506	96.7	522	6	ABK68973	Abk68973 cDNA enco
24	506	96.7	522	6	ABK68970	Abk68970 cDNA enco
25	504.4	96.4	522	6	ABK68966	Abk68966 cDNA enco
26	504.4	96.4	522	6	ABK68968	Abk68968 cDNA enco
27	504.4	96.4	522	6	ABK68967	Abk68967 cDNA enco
28	504.2	96.4	519	3	AAA71313	Aaa71313 Mammalian
29	504.2	96.4	522	3	AAA71315	Aaa71315 Mammalian
30	503.8	96.3	546	2	AAQ97206	Aaq97206 pMON13499
31	503.8	96.3	546	2	AAA03772	Aaa03772 Human G-C
32	503.8	96.3	546	6	ABX00087	Abx00087 Human int
33	503.8	96.3	921	2	AAQ97186	Aaq97186 pMON13026
34	503.8	96.3	921	2	AAQ97189	Aaq97189 pMON13043
35	503.8	96.3	921	2	AAQ97192	Aaq97192 pMON13151
36	503.8	96.3	921	3	AAA03746	Aaa03746 Human int
37	503.8	96.3	921	3	AAA03740	Aaa03740 Human int
38	503.8	96.3	921	3	AAA03743	Aaa03743 Human int
39	503.8	96.3	921	6	ABX00029	Abx00029 Human int
40	503.8	96.3	921	6	ABX00035	Abx00035 Human int
41	503.8	96.3	921	6	ABX00032	Abx00032 Human int
42	503.8	96.3	966	2	AAQ97191	Aaq97191 pMON13045
43	503.8	96.3	966	2	AAQ97187	Aaq97187 pMON13063
44	503.8	96.3	966	2	AAQ97193	Aaq97193 pMON13152
45	503.8	96.3	966	3	AAA03741	Aaa03741 Human int

ALIGNMENTS

RESULT 1
AAN80947
ID   AAN80947 standard; DNA; 525 BP.
XX
AC   AAN80947;
XX
DT   15-NOV-1990   (first entry)
XX
DE   G-CSF gene isolated from peripheral blood macrophages.
XX
KW   Granulocyte colony stimulating factor; hG-CSF; macrophage;
KW   tumour therapy; leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   EP272703-A.
XX
PD   29-JUN-1988.
XX
PF   23-DEC-1987;   87EP-00119157.
XX
PR   23-DEC-1986;   86JP-00306799.
XX
PA   (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI   Kuga T, Komatsu Y, Miyaji H, Sato M, Okabe M, Morimoto M;
PI   Itoh S, Yamasaki M;
XX
DR   WPI; 1988-176825/26.
DR   P-PSDB; AAP80965.
XX
PT   Human granulocyte colony stimulating factor polypeptide derivs. - having
PT   at least one different aminoacid, giving high specific activity and
PT   stability.
XX
PS   Disclosure; Page ?; 68pp; English.
XX
CC   The sequence is carried on plasmid pCSF1-2 isolated from a cDNA library
CC   prepared from mRNA extracted from peripheral blood macro- phages. It
CC   encodes a protein whose AA sequence agrees with those determined for G-
CC   CSF from the human squamous cell line CHU-II and the human bladder cancer
CC   cell line 5637. It can be used as a "master gene" for the construction of
CC   mutant genes which encode variants of hG-CSF which differ by at least one
CC   AA. See also AAN80945 and AAN80946
XX
SQ   Sequence 525 BP; 85 A; 192 C; 151 G; 97 T; 0 U; 0 Other;

```
XX
DR   WPI; 2001-616523/71.
DR   P-PSDB; AAE12154.
XX
PT   Recombinant plasmid vector comprising an endoxylanase signal sequence,
PT   human granulocyte colony stimulating factor gene and other components,
PT   when transformed into microorganism useful for preparing the colony
PT   stimulating factor.
XX
PS   Example 2; Fig 5; 50pp; English.
XX
CC   The invention relates to an Escherichia coli producing and secreting
CC   human granulocyte colony stimulating factor (hG-CSF), more specifically,
CC   to a recombinant plasmid constructed to express secretory hG-CSF in E.
CC   coli, an E. coli transformed with that plasmid to secrete hG-CSF, and a
CC   process for preparing hG-CSF using the transformed hG-CSF. The
CC   recombinant plasmid vector comprises of a kanamycin resistance gene, a
CC   promoter, an endoxylanase signal sequence, a nucleotide sequence coding
CC   for an oligopeptide consisting of 13 amino acids including 6 consecutive
CC   histidine residues and a hG-CSF. E.coli transformed with recombinant
CC   plasmid vector is useful for preparing hG-CSF. The method comprises
CC   culturing the microorganism to obtain a hG-CSF fusion protein and
CC   treating the fusion protein with a protease preferably Factor Xa, to
CC   obtain a hG-CSF, where the fusion protein is obtained from the culture by
CC   employing Ni-column. The present sequence is human hG-CSF gene inserted
CC   into plasmid pEDCSFm.
XX
SQ   Sequence 531 BP; 89 A; 192 C; 151 G; 99 T; 0 U; 0 Other;

Query Match             100.0%;  Score 523;  DB 4;  Length 531;
Best Local Similarity   100.0%;  Pred. No. 2.2e-107;
Matches   523;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ACCCCCTGGGCCCTGCCCAGCTCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 ACCCCCTGGGCCCTGCCCAGCTCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 63

Qy         61 GTGAGGAAGATCAGGGCCGATGGCGGCAGGCTCCAGGAGAGAGCTGTGTGCCACCTACAAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GTGAGGAAGATCAGGGCCGATGGCGGCAGGCTCCAGGAGAGAGCTGTGTGCCACCTACAAG 123

Qy        121 CTGTGCCACCCGGAGAGCTGGTGCTCGGACACTCTCTGGGCATCCCTCTGGGCTCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CTGTGCCACCCGGAGAGCTGGTGCTCGGACACTCTCTGGGCATCCCTCTGGGCTCCC 183

Qy        181 CTGAGCAGCTGCCCAGCCAGCCCTGGCAGGCCCTGCTTGAGCCAACTCCATAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 CTGAGCAGCTGCCCAGCCAGCCCTGCGAGCTGCCCAGCCCTGCTTGAGCCAACTCCATAGC 243

Qy        241 GGCCTTTCCTCTACCAGGGGCTCGCGCCGCCTCCGGAAGGGATCTCCCGGAGGTTGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GGCCTTTCCTCTACCAGGGGCTCGCGCCGCCTCCGGAAGGGATCTCCCGGAGGTTGGGT 303

Qy        301 CCACCTTGGACACACTGCAGTCGAGCGTGCGCCGACTTTGCCACCATCTGGCAGCAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 CCACCTTGGACACACTGCAGTCGAGCGTGCGCCGACTTTGCCACCATCTGGCAGCAG 363

Qy        361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 423

Qy        421 GCCTCTGCTTTCCAGCGCGCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GCCTCTGCTTTCCAGCGCGCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 483

Qy        481 CTGGAGGTGTGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
              |||||||||||||||||||||||||||||||||||||||||||
Db        484 CTGGAGGTGTGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 526

RESULT 3
AAD19771
ID   AAD19771 standard; DNA; 615 BP.
```

```
Query Match             100.0%;  Score 523;  DB 1;  Length 525;
Best Local Similarity   100.0%;  Pred. No. 2.2e-107;
Matches   523;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ACCCCCTGGGCCCTGCCCAGCTCCTGCCCCCAGAGCTTCTGCTCAAGTGCTTAGAGCAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCCCCTGGGCCCTGCCCAGCTCCTGCCCCCAGAGCTTCTGCTCAAGTGCTTAGAGCAA 60

Qy         61 GTGAGGAAGATCAGGGCCATGGCGGCAGGCTCCAGGAGAGAGCTGTGTGCCACCTACAAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGAGGAAGATCAGGGCCATGGCGGCAGGCTCCAGGAGAGAGCTGTGTGCCACCTACAAG 120

Qy        121 CTGTGCCACCCGGAGGAGTGGTGCTCGGACACTCTCTGGGCTCCC 180
              |||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGTGCCACCCGGAGGAGTGGTGCTCGGACACTCTCTGGGCTCCC 180

Qy        181 CTGAGCAGCTGCCCAGCCAGCCCTGGCAGCTGCTTGAGCCAACTCCATAGC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGCAGCTGCCCAGCCAGCCCTGGCAGCTGCTTGAGCCAACTCCATAGC 240

Qy        241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGGCCCTGGAAGGGATCTCCCGAGTTGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGGCCCTGGAAGGGATCTCCCGAGTTGGGT 300

Qy        301 CCACCTTGGACACACTGCAGTCGACGTGCGCCGACTTTGCCACCATCTGGCAGCAG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCACCTTGGACACACTGCAGTCGACGTGCGCCGACTTTGCCACCATCTGGCAGCAG 360

Qy        361 ATGGAAGAACTGGGAATGGCCCCGGCAGGAGGGGTCCATGCCGGCCTTC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGAAGAACTGGGAATGGCCCCGGCAGGAGGGGTCCATGCCGGCCTTC 420

Qy        421 GCCTCTGCTTTCCAGCGCGCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCTCTGCTTTCCAGCGCGCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480

Qy        481 CTGGAGGTGTGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
              |||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGGAGGTGTGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523

RESULT 2
AAD19772
ID   AAD19772 standard; DNA; 531 BP.
XX
AC   AAD19772;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human hG-CSF gene inserted into plasmid pEDCSFm.
XX
KW   Human; granulocyte colony stimulating factor; hG-CSF; protease;
KW   Factor Xa; kanamycin resistance; endoxylanase signal peptide; ds.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..528
FT                  /*tag= a
FT                  /product= "Human hG-CSF protein"
XX
PN   WO200173081-A1.
XX
PD   04-OCT-2001.
XX
PF   31-MAR-2001; 2001WO-KR000549.
XX
PR   31-MAR-2000; 2000KR-00017052.
XX
PA   (KOAD ) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI   Lee S, Jeong K;
```

```
XX
AC   AAD19771;
XX
DT   18-DEC-2001 (first entry)
XX
DE   Human hG-CSF gene inserted into plasmid p19CSFm.
XX
KW   Human; granulocyte colony stimulating factor; hG-CSF; protease;
KW   Factor Xa; kanamycin resistance; endoxylanase signal peptide; ds.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            88..612
FT                  /tag= a
FT                  /product= "Human hG-CSF protein"
FT                  /note= "CDS does not include start codon"
FT                  /partial
XX
PN   WO200173081-A1.
XX
PD   04-OCT-2001.
XX
PF   31-MAR-2001; 2001WO-KR000549.
XX
PR   31-MAR-2000; 2000KR-00017052.
XX
PA   (KOAD ) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI   Lee S, Jeong K;
XX
DR   WPI; 2001-616523/71.
DR   P-PSDB; AAE12153.
XX
PT   Recombinant plasmid vector comprising an endoxylanase signal sequence,
PT   human granulocyte colony stimulating factor gene and other components,
PT   when transformed into microorganism useful for preparing the colony
PT   stimulating factor.
XX
PS   Example 1; Fig 3; 50pp; English.
XX
CC   The invention relates to an Escherichia coli producing and secreting
CC   human granulocyte colony stimulating factor (hG-CSF), more specifically,
CC   to a recombinant plasmid constructed to express secretory hG-CSF in E.
CC   coli, an E. coli transformed with that plasmid to secrete hG-CSF, and a
CC   process for preparing hG-CSF using the transformed hG-CSF. The
CC   recombinant plasmid vector comprises of a kanamycin resistance gene, a
CC   promoter, an endoxylanase signal sequence, a nucleotide sequence coding
CC   for an oligopeptide consisting of 13 amino acids including 6 consecutive
CC   histidine residues and a hG-CSF. E.coli transformed with recombinant
CC   plasmid vector is useful for preparing hG-CSF. The method comprises
CC   culturing the microorganism to obtain a hG-CSF fusion protein and
CC   treating the fusion protein with a protease preferably Factor Xa, to
CC   obtain a hG-CSF, where the fusion protein is obtained from the culture by
CC   employing Ni-column. The present sequence is human hG-CSF gene inserted
CC   into plasmid p19CSFm
XX
SQ   Sequence 615 BP; 105 A; 219 C; 178 G; 113 T; 0 U; 0 Other;

Query Match      100.0%;  Score 523;  DB 4;  Length 615;
Best Local Similarity   100.0%;  Pred. No. 2.2e-107;
Matches 523;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACCCCCCTGGCCCTGCCCCCAGCTTCCTGCTCCAGCTTCTGGCTCAAGTGCTTAGAGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   88 ACCCCCCTGGCCCTGCCCCCAGCTTCCTGCTCCAGCTTCTGGCTCAAGTGCTTAGAGCAA 147

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCAGCGGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  148 GTGAGGAAGATCCAGGGCGATGGGCAGCGGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 207

Qy  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  208 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 267

Qy  181 CTGAGCAGTGCCCCAGCCCAGCTGGCAGGCCCTGCTGCAGGCTGCCAGCCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  268 CTGAGCAGTGCCCCAGCCCAGCTGGCAGGCCCTGCTGCAGGCTGCCAGCCAACTCCATAGC 327

Qy  241 GGCCTTTCCTTACCAGGGGATCTCCAGGGCCCTGGAAGGGGATCTCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  328 GGCCTTTCCTTACCAGGGGATCTCCAGGGCCCTGGAAGGGGATCTCCCCGAGTTGGGT 387

Qy  301 CCCACCTTGGACACACTGCAGCTGGACGTGCGCCGACTTGCCACCACCATCTGGCAGCAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  388 CCCACCTTGGACACACTGCAGCTGGACGTGCGCCGACTTGCCACCACCATCTGGCAGCAG 447

Qy  361 ATGGAAGAACTGGAATGCCCCTGCCCTGCAGCCCTGCAGCCCCAGGGTGCCATGGCCGGCCTTC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  448 ATGGAAGAACTGGAATGCCCCTGCCCTGCAGCCCTGCAGCCCCAGGGTGCCATGGCCGGCCTTC 507

Qy  421 GCCTCTGCTTCCAGCGCCGCGGAGGACGGGGTCCTAGTTGCCTCCATCTGCAGAGCTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  508 GCCTCTGCTTCCAGCGCCGCGGAGGACGGGGTCCTAGTTGCCTCCATCTGCAGAGCTTC 567

Qy  481 CTGGAGGTGTCGTACCGGCTTCTACGCCACCTTGCCCAGCCCT 523
         |||||||||||||||||||||||||||||||||||||||||||
Db  568 CTGGAGGTGTCGTACCGGCTTCTACGCCACCTTGCCCAGCCCT 610
```

```
RESULT 4
AAN70223
ID   AAN70223 standard; DNA; 1520 BP.
XX
AC   AAN70223;
XX
DT   25-MAR-2003 (revised)
DT   25-APR-1991 (first entry)
XX
DE   Plasmid pBRV2 insert.
XX
KW   Human granulocyte colony stimulating factor; G-CSF; leukaemia; ss.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            31..645
FT                  /tag= a
XX
PN   EP220520-A.
XX
PD   06-MAY-1987.
XX
PF   30-SEP-1986;   86EP-00113446.
XX
PR   17-SEP-1985;   85JP-00206066.
PR   20-SEP-1985;   85JP-00209638.
PR   30-SEP-1985;   85JP-00217150.
PR   02-DEC-1985;   85JP-00269455.
PR   02-DEC-1985;   85JP-00269456.
PR   03-DEC-1985;   85JP-00270838.
PR   03-DEC-1985;   85JP-00270839.
PR   17-JUL-1986;   86JP-00166709.
PR   17-JUL-1986;   86JP-00166710.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
XX
PI   Yamazaki T, Nagata S, Tsuchiya M;
XX
DR   WPI; 1987-124182/18.
DR   P-PSDB; AAP70162.
XX
PT   Poly:peptide with human granulocyte colony stimulating factor activity -
PT   is obtd. by cultivating transformant formed by recombinant DNA
PT   procedures.
XX
```

PN   EP217404-A.
XX
PD   08-APR-1987.
XX
PF   03-OCT-1986; 86EP-00113671.
XX
PR   04-OCT-1985; 85JP-00220450.
PR   02-JUN-1986; 86JP-00125660.
XX
PA   (CHUS) CHUGAI SEIYAKU KK.
XX
PI   Tamura M, Nomura H, Hattori K, Ono M;
XX
DR   WPI; 1987-095431/14.
DR   P-PSDB; AAP71384.
XX
PT   Leukopenia treating agent esp. for producing fully mature neutrophiles -
PT   contains human granulocyte colony stimulating factor obtd. by recombinant
PT   DNA methods etc.
XX
PS   Claim 7; Fig 2; 34pp; English.
XX
CC   The hGCSF has the following properties: (i) mol. wgt. 19000 +/- 1000 (by
CC   SDS-PAGE); (ii) isoelectric pt. at 5.5, 5.8 or 6.1 each +/- 0.1; (iii) UV
CC   max. absorption at 280nm and min. at 250nm. Prodn. of the gene, vector
CC   etc. is described in JP 269455, 269456, 270838 and 270839, each of 1985
XX
SQ   Sequence 1520 BP; 305 A; 488 C; 401 G; 326 T; 0 U; 0 Other;

Query Match   100.0%; Score 523; DB 1; Length 1520;
Best Local Similarity   100.0%; Pred. No. 2.5e-107;
Matches   523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 180

Qy   61 GTGAGGAAGATCCAGGGCGCAGCGGCCGATGGGCGCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGAGGAAGATCCAGGGCGCAGCGGCCGATGGGCGCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAG 240

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGTGCCACCCCGAGGAGCTGGTGCTGTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 300

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCGTCGTTGAGCCAACTCCATAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCGTCGTTGAGCCAACTCCATAGC 360

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCGCGAGGGGCTCGCGAGGGATCTCCCCGAGTTGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCCTTTTCCTCTACCAGGGGCTCGCGAGGGGCTCGCGAGGGATCTCCCCGAGTTGGGT 420

Qy   301 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTGCCACCACCATCCGCAGCAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTGCCACCACCATCCGCAGCAG 480

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCAGCCCACCCAGGGTGCCATGCCGCGGCCTTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGGAAGAACTGGGAATGGCCCCTGCAGCCCACCCAGGGTGCCATGCCGCGGCCTTC 540

Qy   421 GCCTCTGCTTTCGTACGGCCGTTCTCCATCTGCCTCCCATCTGCAGAGCTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCCTCTGCTTTCGTACGGCCGTTCTCCATCTGCCTCCCATCTGCAGAGCTTC 600

Qy   481 CTGGAGGTGTGTACCGCGTTCACGCCACCTTGCCCAGCCCT 523
              ||||||||||||||||||||||||||||||||||||||||
Db   601 CTGGAGGTGTGTACCGCGTTCACGCCACCTTGCCCAGCCCT 643

RESULT 6
AAN60937
ID   AAN60937 standard; DNA; 1521 BP.

PS   Disclosure; Fig 4; 73pp; English.
XX
CC   The plasmid was isolated from a cDNA library prepd. from CHU-2 cells, a
CC   human oral cavity tumour cell line, using a 1500 bp insert from pBRG4
CC   (AAN70222) and probe LC (AAN13351. The plasmid was used to prepare
CC   recombinant expression plasmids for the prodn. of h G-CSF. See also
CC   AAN70221-N70224 and AAN71349-N71351. (Updated on 25-MAR-2003 to correct
CC   PR field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 1520 BP; 303 A; 487 C; 402 G; 328 T; 0 U; 0 Other;

Query Match   100.0%; Score 523; DB 1; Length 1520;
Best Local Similarity   100.0%; Pred. No. 2.5e-107;
Matches   523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 180

Qy   61 GTGAGGAAGATCCAGGGCGCAGCGGCCGATGGGCGCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGAGGAAGATCCAGGGCGCAGCGGCCGATGGGCGCAGCGCTCCAGGAGAGCTGTGTGCCACCTACAAG 240

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGTGCCACCCCGAGGAGCTGGTGCTGTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 300

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCGTCGTTGAGCCAACTCCATAGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGCAGGCGTCGTTGAGCCAACTCCATAGC 360

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCGCGAGGGGCTCGCGAGGGATCTCCCCGAGTTGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCCTTTTCCTCTACCAGGGGCTCGCGAGGGGCTCGCGAGGGATCTCCCCGAGTTGGGT 420

Qy   301 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTGCCACCACCATCCGCAGCAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTGCCACCACCATCCGCAGCAG 480

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCAGCCCACCCAGGGTGCCATGCCGCGGCCTTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGGAAGAACTGGGAATGGCCCCTGCAGCCCACCCAGGGTGCCATGCCGCGGCCTTC 540

Qy   421 GCCTCTGCTTTCGTACGGCCGTTCTCCATCTGCCTCCCATCTGCAGAGCTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCCTCTGCTTTCGTACGGCCGTTCTCCATCTGCCTCCCATCTGCAGAGCTTC 600

Qy   481 CTGGAGGTGTGTACCGCGTTCACGCCACCTTGCCCAGCCCT 523
              ||||||||||||||||||||||||||||||||||||||||
Db   601 CTGGAGGTGTGTACCGCGTTCACGCCACCTTGCCCAGCCCT 643

RESULT 5
AAN71320
ID   AAN71320 standard; cDNA; 1520 BP.
XX
AC   AAN71320;
XX
DT   23-APR-1991 (first entry)
XX
DE   Sequence encoding human granulocyte colony stimulating factor (hGCSF) in
DE   pBRV2.
XX
KW   Leukopenia therapy; neutropenia; eosinopenia; lymphopenia; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   31..120
FT   /*tag= a
FT   mat_peptide   121..645
FT   /*tag= b
XX

```
Db    421 CCCACCTTGGACACACTGCAGCTGGACGTGCGCCGACTTTGCCACCACCATCTGGCAGCAG 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 ATGGAAGAACTGGGAATGGCCCTGCCCTGCGCCCCAGGGTGCCATGCCCGGCCTTC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGGAAGAACTGGGAATGGCCCTGCCCTGCGCCCCAGGGTGCCATGCCCGGCCTTC 540

Qy    421 GCCCTCGCTTCCAGCGCCGGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCCCTCGCTTCCAGCGCCGGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 600

Qy    481 CTGGAGGTGTCGTACCGGCGTTCTACCGCCACCTTGCCCAGCCCT 523
                |||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGGAGGTGTCGTACCGGCGTTCTACCGCCACCTTGCCCAGCCCT 643
```

```
RESULT 7
AAN81478
ID    AAN81478 standard; cDNA; 1525 BP.
XX
AC    AAN81478;
XX
DT    25-MAR-2003 (revised)
DT    14-DEC-1990 (first entry)
XX
DE    Sequence encoding human granulocyte colony-stimulating factor (G-CSF) in
DE    clone pPl2 of MIA PaCa cDNA library.
XX
KW    Leukaemia therapy; anaemia treatment; leukocyte; lymphokine;
KW    interleukin-3; ss.
XX
OS    Homo sapiens.
XX
FH    Key              Location/Qualifiers
FT
FT    CDS              43..132
FT                     /*tag= a
FT    mat_peptide      133..657
FT                     /*tag= b
FT    conflict         237..238
FT                     /*tag= c
FT                     /note= "CHU-2 cDNA clones of Nagata et al. contains a 9
FT                     base pair insertion (GTGATGGAG)"
FT    conflict         588
FT                     /*tag= d
FT                     /note= "G in the CHU-2 clone"
FT    conflict         1237
FT                     /*tag= e
FT                     /note= "C in the CHU-2 clone"
XX
PN    EP255843-A.
XX
PD    24-FEB-1988.
XX
PF    11-AUG-1987;    87EP-00307114.
XX
PR    11-AUG-1986;    86US-00895154.
PR    18-NOV-1986;    86US-00932037.
XX
PA    (CETU ) CETUS CORP.
XX
PI    Devlin JJ, Devlin PE, Kawaskai ES,   Warren MK;
XX
DR    WPI; 1988-051585/08.
DR    P-PSDB; AAP1162.
XX
PT    Human recombinant granulocyte colony-stimulating factor - causes
PT    differentiation of granulocytes and monocytes from bone marrow progenitor
PT    cells.
XX
PS    Example; Fig 4; 61pp; English.
XX
CC    The major difference between this clone and the CHU-2 cDNA clones of
CC    Nagata et al. (EMBO Journal 5:3,575-581(1986)) are given in FT of
CC    AAN81478. The G-CSF was detected in purified mRNA fraction obtd. by
```

```
XX
AC    AAN60937;
XX
DT    18-OCT-1991 (first entry)
XX
DE    Plasmid pBRV2 insert encoding granulocyte CSF.
XX
KW    G-CSF; granulocyte colony stimulating factor; antimicrobial.
XX
OS    Synthetic.
XX
FH    Key              Location/Qualifiers
FT
FT    CDS              31..645
FT                     /*tag= a
FT    mat_peptide      121
FT                     /*tag= b
XX
PN    WO8604506-A.
XX
PD    14-AUG-1986.
XX
PF    07-FEB-1986;    86WO-JP000053.
XX
PR    08-FEB-1985;    85JP-00023777.
PR    03-DEC-1985;    85JP-00270839.
XX
PA    (CHUS ) CHUGAI SEIYAKU KK.
PA    (ONOM/) ONO M.
XX
PI    Ono M, Nomura H, Tamura M, Matsumoto M;
XX
DR    WPI; 1986-225384/34.
DR    P-PSDB; AAP61341.
XX
PT    Infection-protection inducing agent - contains human granulocyte colony
PT    stimulating factor obtd. by cell culture or recombinant technology.
XX
PS    Disclosure; Fig 4; 139pp; Japanese.
XX
CC    The plasmid encodes a novel polypeptide having human graulocyte colony
CC    stimulating factor activity. The CSF may be produced either from a cell-
CC    line secreting the factor into its medium, or from a transformant
CC    microorganism. The product induces high immunity to infection by
CC    anaerobic and aerobic bacteria and fungi
XX
SQ    Sequence 1521 BP; 304 A; 488 C; 402 G; 327 T; 0 U; 0 Other;

Query Match        100.0%;  Score 523; DB 1; Length 1521;
Best Local Similarity  100.0%;  Pred. No. 2.5e-107;
Matches 523;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy      1 ACCCCCCTGGGCCCTGCCAGCTCCGCAGCTTCCTGCTTCAAGTGCTTAGAGCAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACCCCCCTGGGCCCTGCCAGCTCCGCAGCTTCCTGCTTCAAGTGCTTAGAGCAA 180

Qy     61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 240

Qy    121 CTGTGCCACCCCGAGGAGAGTCGGTGCTGCTTCGGGACACTCCTGGGCTCCC 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGTGCCACCCCGAGGAGAGTCGGTGCTGCTTCGGGACACTCCTGGGCTCCC 300

Qy    181 CTGAGCAGCTGCCCCAGGCCCTGCAGCTGGCAGGCCCTGGCTTGAGCCAACTCCATAGC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGAGCAGCTGCCCCAGGCCCTGCAGCTGGCAGGCCCTGGCTTGAGCCAACTCCATAGC 360

Qy    241 GGCCTTTTCCTCTACCAGGGGCTCCTGCCAGGCCGTCCGGAAGGGATCTCCCCCGAGTTGGGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCCTTTTCCTCTACCAGGGGCTCCTGCCAGGCCGTCCGGAAGGGATCTCCCCCGAGTTGGGT 420

Qy    301 CCCACCTTGGACACACTGCAGCTGCGCCGACTTTGCCACCACCATCTGGCAGCAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 8
AAN91086
```
ID   AAN91086 standard; DNA; 1525 BP.
XX
AC   AAN91086;
XX
DT   04-JUL-1990  (first entry)
XX
DE   Plasmid pP12 contg. colony stimulating factor-1 gene.
XX
KW   Colony stimulating factor-1; pP12; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     43..132
FT                   /*tag= a
FT   mat_peptide     133..654
FT                   /*tag= b
XX
PN   WO8901038-A.
XX
PD   09-FEB-1989.
XX
PF   20-JUL-1988;  88WO-US002445.
XX
PR   24-JUL-1987;  87US-00077188.
XX
PA   (CETU ) CETUS CORP.
XX
PI   Kawasaki ES, Devlin JJ, Martin G, O'rouke E, Clark R;
XX
DR   WPI; 1989-061174/08.
DR   P-PSDB; AAP95033.
XX
PT   Recombinant Baculovirus transfer vectors - used for prodn. of colony
PT   stimulating factor in Baculovirus-insect cell expression system.
XX
PS   Disclosure; Fig 3; 68pp; English.
XX
CC   The cDNA insert of pP12 contains 11 more bases than the CHU-2 G-CSF
CC   clone. The major difference between this clone, derived from MIA PaCa-2,
CC   and CHU-2 clone is a 9 bp insertion in the CHU-2 clone between bases 237
CC   and 238. There are two other differences: an A at position 588 (G in the
CC   CHU-2 clone) is a silent third base change, and a T at position 1237(C in
CC   the CHU-2 clone) in the 3' untranslated region. See also AAN91085
CC   sucrose density gradient centrifugation of induced MIA PaCa-2 cells as
CC   described in WO804607. (Updated on 25-MAR-2003 to correct PR field.)
CC   (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 1525 BP; 303 A; 491 C; 402 G; 329 T; 0 U; 0 Other;
```

```
Query Match             100.0%;  Score 523;  DB 1;  Length 1525;
Best Local Similarity   100.0%;  Pred. No. 2.5e-107;
Matches  523;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 ACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133 ACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 192

Qy    61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 252

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTTCTGGGCATCCCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTTCTGGGCATCCCTGGGCTCCC 312

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGCAGGCGGCCCTGCTTGAGCCAACTCCATAGC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGCAGGCGGCCCTGCTTGAGCCAACTCCATAGC 372

Qy   241 GGCCTTTCCTCTACCAGGGGATCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   373 GGCCTTTCCTCTACCAGGGGATCCCTGGAAGGGATCTCCCCGAGTTGGGT 432

Qy   301 CCCACCTTGGACACACTGACGTGGAATGGCCCCTGCCAGCGCCCTGCCATGGCCGGCCAG 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433 CCCACCTTGGACACACTGACGTGGAATGGCCCCTGCCAGCGCCCTGCCATGGCCGGCCAG 492

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCAGCCCCAGGGGTGCCATGGCCGGCCTTC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCAGCCCCAGGGGTGCCATGGCCGGCCTTC 552

Qy   421 GCCTCTGCTTTCCAGCGCCGGCAGGAGGGGTCCTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   553 GCCTCTGCTTTCCAGCGCCGGCAGGAGGGGTCCTCCTAGTTGCCTCCCATCTGCAGAGCTTC 612

Qy   481 CTGGAGGTGTCGTACCGTTCTACGGCCACCTTGCCCAGCCCT 523
         ||||||||||||||||||||||||||||||||||||||||||
Db   613 CTGGAGGTGTCGTACCGTTCTACGGCCACCTTGCCCAGCCCT 655
```

```
SQ   Sequence 1525 BP; 303 A; 490 C; 403 G; 328 T; 0 U; 1 Other;
```

```
Query Match             100.0%;  Score 523;  DB 1;  Length 1525;
Best Local Similarity   100.0%;  Pred. No. 2.5e-107;
Matches  523;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 ACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133 ACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 192

Qy    61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193 GTGAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 252

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTTCTGGGCATCCCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTTCTGGGCATCCCTGGGCTCCC 312

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGCAGGCGGCCCTGCTTGAGCCAACTCCATAGC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGCAGGCGGCCCTGCTTGAGCCAACTCCATAGC 372

Qy   241 GGCCTTTCCTCTACCAGGGGATCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   373 GGCCTTTCCTCTACCAGGGGATCCCTGGAAGGGATCTCCCCGAGTTGGGT 432

Qy   301 CCCACCTTGGACACACTGACGTGGAATGGCCCCTGCCAGCGCCCTGCCATGGCCGGCCTTC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433 CCCACCTTGGACACACTGACGTGGAATGGCCCCTGCCAGCGCCCTGCCATGGCCGGCCTTC 492

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCAGCCCCAGGGGTGCCATGGCCGGCCTTC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCAGCCCCAGGGGTGCCATGGCCGGCCTTC 552

Qy   421 GCCTCTGCTTTCCAGCGCCGGCAGGAGGGGTCCTCCCATCTGCAGAGCTTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   553 GCCTCTGCTTTCCAGCGCCGGCAGGAGGGGTCCTCCCATCTGCAGAGCTTC 612

Qy   481 CTGGAGGTGTCGTACCGTTCTACGGCCACCTTGCCCAGCCCT 523
         ||||||||||||||||||||||||||||||||||||||||||
Db   613 CTGGAGGTGTCGTACCGTTCTACGGCCACCTTGCCCAGCCCT 655
```

RESULT 9
ACC78869
```
ID   ACC78869 standard; DNA; 525 BP.
XX
AC   ACC78869;
XX
DT   02-SEP-2003  (first entry)
XX
DE   Human granulocyte colony stimulating factor (G-CSF) encoding DNA.
```

```
Db       1 ACCCCCTGGGCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy      61 GTGAGGAAGATCCAGGGCGATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      61 GTGAGGAAGATCCAGGGCGATGGCGCAGGCGCTGCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy     121 CTGTGCCACCCCGAGGAGCTGTGTCTGCTGCACGACACTCTCTGGGCATCCCCTGGGCTCCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTGTGCCACCCCGAGGAGCTGTGTCTGCTGCACGACACTCTCTGGGCATCCCCTGGGCTCCC 180

Qy     181 CTGAGCAGTGCCCCAGCCCAGCCAGGCCCTGCCAGCTGCGTCTTGAGCCAACTCCATAGC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGAGCAGTGCCCCAGCCCAGCCAGGCCCTGCCAGCTGCGTCTTGAGCCAACTCCATAGC 240

Qy     241 GGCCTTTCCTTACCAGGGGCCTCCTGGAGGGGATCTCCCCGAGTTGGGT 300
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGCCTTTCCTTCCTTACCAGGGGCCTCCTGGAGGGGATCTCCCCGAGTTGGGT 300

Qy     301 CCCACCTTGGACACACTGCAGCTGGACGTGCGCGACTTTGCCACCATCTGGCAGCAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CCCACCTTGGACACACTGCAGTTGGACGTGCGCGACTTTGCCACCATCTGGCAGCAG 360

Qy     361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 420

Qy     421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GCCTCTCGCTTTCCAGCGCCGGGCAGGAGGGGTCCTGGTTGCCTCCCATCTGCAGAGCTTC 480

Qy     481 CTGGAGGTGTCGTACCGGGTTCTACGCCACCTTGCCCAGCCCT 523
           |||||||||||||||||||||||||||||||||||||||||||
Db     481 CTGGAGGTGTCGTACCGGGTTCTACGCCACCTTGCCCAGCCCT 523
```

RESULT 10
ABA94177
ID   ABA94177 standard; cDNA; 549 BP.
XX
AC   ABA94177;
XX
DT   08-MAY-2002 (first entry)
XX
DE   Granulocyte colony stimulating factor related nucleotide sequence.
XX
KW   Granulocyte colony stimulating factor; G-CSF; gene; ss.
XX
OS   Unidentified.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..549
FT   /*tag= a
FT   /partial
FT   /product= "G-CSF"
FT   /note= "no start codon given"
XX
PN   KR98077885-A.
XX
PD   16-NOV-1998.
XX
PF   23-APR-1997; 97KR-00015210.
XX
PR   23-APR-1997; 97KR-00015210.
XX
PA   (MOKA-) MOKAM LIFE SCI RES CENT.
XX
PI   Sohn YD, Lee EG, Kim SH, Park DH;
XX
DR   WPI; 2000-035728/03.
DR   P-PSDB; ABB05960.
XX
PT   Recombinant microorganism expressing granulocyte colony stimulating
PT   factor and process for preparing the recombinant protein therefrom.

XX
KW   Peptide remodeling; glycoconjugation; glycosyltransferase; glycan;
KW   granulocyte colony stimulating factor; G-CSF; human; gene; ds.
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..525
FT   /*tag= a
FT   /product= "G-CSF"
XX
PN   WO200031464-A2.
XX
PD   17-APR-2003.
XX
PF   09-OCT-2002; 2002WO-US032263.
XX
PR   10-OCT-2001; 2001US-0328523P.
PR   19-OCT-2001; 2001US-0344692P.
PR   28-NOV-2001; 2001US-0334233P.
PR   28-NOV-2001; 2001US-0334301P.
PR   07-JUN-2002; 2002US-0387292P.
PR   25-JUN-2002; 2002US-0391777P.
PR   17-JUL-2002; 2002US-0396594P.
PR   16-AUG-2002; 2002US-0404249P.
PR   28-AUG-2002; 2002US-0407527P.
XX
PA   (NEOS-) NEOSE TECHNOLOGIES INC.
XX
PI   De Frees S, Zopf D, Bayer R, Bowe C, Hakes D, Chen X;
XX
DR   WPI; 2003-449162/42.
DR   P-PSDB; ABR55839.
XX
PT   Remodeling a peptide, by removing a saccharyl subunit from the peptide to
PT   form truncated glycan, and adding or deleting glycosyl groups to a
PT   peptide and/or adding modifying group of a peptide to remodel the
PT   peptide.
XX
PS   Example; Fig 52; 900pp; English.
XX
CC   The invention relates to a cell-free, in vitro method of remodeling a
CC   peptide. The method involves removing a saccharyl subunit from the
CC   peptide, thus forming a truncated glycan, and contacting the truncated
CC   glycan with at least one glycosyltransferase and at least one glycosyl
CC   donor under conditions suitable to transfer at least one glycosyl donor
CC   to the truncated glycan, thus remodeling the peptide. Conjugates can be
CC   formed between a granulocyte colony stimulating factor (G-CSF) peptide,
CC   interferon alpha peptide, interferon beta peptide, Factor VIIa peptide,
CC   Factor IX peptide, follicle stimulating hormone peptide, erythropoietin
CC   (EPO) peptide, granulocyte macrophage colony stimulating factor (GM-CSF)
CC   peptide, interferon-gamma peptide, alpha-1-protease inhibitor (A-1-PI)
CC   peptide, beta-glucosidase peptide, tissue plasminogen activator (TPA)
CC   peptide, interleukin-2 (IL-2) peptide, Factor VIII peptide, TNFalpha
CC   receptor/immunoglobulin (Ig) G fusion peptide, urokinase peptide, anti-
CC   glycoprotein IIb/IIIa monoclonal antibody peptide, chimeric anti HER2
CC   antibody peptide, anti-respiratory syncytial virus (RSV) F peptide, anti-
CC   CD20 antibody peptide, recombinant DNase peptide, anti-TNF alpha peptide,
CC   insulin peptide, hepatitis B surface antigen (HbsAg), human growth
CC   hormone (HGH) peptide, and a modifying group, where the modifying group
CC   is covalently attached to the peptide through an intact glycosyl linking
CC   group. The method is useful for a cell-free, in vitro method of
CC   remodeling the above mentioned peptides. The present sequence represents
CC   a human G-CSF encoding DNA
XX
SQ   Sequence 525 BP; 84 A; 192 C; 152 G; 97 T; 0 U; 0 Other;

```
Query Match    99.7%;  Score 521.4;  DB 7;  Length 525;
Best Local Similarity  99.8%;  Pred. No. 4.9e-107;
Matches 522;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy       1 ACCCCCTGGGCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

XX
PS Disclosure; Page 11; 16pp; Korean.
XX
CC The present invention describes a recombinant microorganism expressing a
CC granulocyte colony stimulating factor (G-CSF), and process for preparing
CC the recombinant protein. The present sequence encodes a G-CSF related
CC protein, which is given in the exemplification of the present invention
XX
SQ Sequence 549 BP; 90 A; 199 C; 160 T; 100 T; 0 U; 0 Other;

Query Match   99.7%; Score 521.4; DB 3; Length 549;
Best Local Similarity   99.8%; Pred. No. 5e-107;
Matches   522; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy    1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 84

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGCAGGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GTGAGGAAGATCCAGGGCGATGGCGCAGGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 144

Qy  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  145 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 204

Qy  181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGCGGCTGCTTGAGCCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGCGGCTGCTTGAGCCAACTCCATAGC 264

Qy  241 GGCCTTTCCTTACCAGGGGTCCTCGGACGCCCTGGAAGGATCCCCGAGGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 GGCCTTTCCTTACCAGGGGTCCTCGGACGCCCTGGAAGGATCCCCGAGGTTGGGT 324

Qy  301 CCCACCTTGGACACACTGCAGTGACGTCGCCGACTTTGCCACCACCATCTGGCAGCAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 CCCACCTTGGACACACTGCAGTGACGTCGCCGACTTTGCCACCACCATCTGGCAGCAG 384

Qy  361 ATGGAAGAACTGGGAATGGCGCGGCGGCAGGAGGGGGTCCATGCCGGCCGGCCTTC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  385 ATGGAAGAACTGGGAATGGCGCGGCGGCAGGAGGGGGTCCATGCCGGCCGGCCTTC 444

Qy  421 GCCTCTGCTTTCAGCGCGGCGGGAGGGGGTCCTCCATCTGCAGAGCTTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  445 GCCTCTGCTTTCAGCGCGGCGGGAGGGGGTCCTCCATCTGCAGAGCTTC 504

Qy  481 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
         |||||||||||||||||||||||||||||||||||||||||||
Db  505 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 547
```

RESULT 11
AAN71089
ID   AAN71089 standard; DNA; 1415 BP.
XX
AC   AAN71089;
XX
DT   25-MAR-2003 (revised)
DT   26-APR-1991 (first entry)
XX
DE   Sequence of human granulocyte colony stimulating factor (hpG-CSF) cDNA
DE   clone Ppo2.
XX
KW   Haematopoietic disorders; therapy; aplastic anaemia;
KW   bone marrow transplant; burn wounds; leukaemia; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..36
FT   /*tag= a
FT   mat_peptide   37..561
FT   /*tag= b

FT   polyA_signal   1797..1802
FT   /*tag= c
XX
PN   WO8701132-A.
PD   26-FEB-1987.
XX
PF   22-AUG-1986; 86WO-US001708.
XX
PR   23-AUG-1985; 85US-00768954.
PR   23-AUG-1985; 85US-00768959.
PR   03-MAR-1986; 86US-00835548.
XX
PA   (KIRI) KIRIN AMGEN INC.
PA   (KIRI) KIRIN AMGEN INC.
PA   (AMGE-) AMGEN.
XX
PI   Souza LM;
XX
DR   WPI; 1987-064855/09.
DR   P-PSDB; AAP70730.
XX
PT   Poly:peptide with granulocyte colony stimulating factor activity - obtd.
PT   by recombinant DNA procedures for treating haematopoietic disorders.
XX
PS   Disclosure; Page 22-24; 79pp; English.
XX
CC   The examples describe procedures for the designing of probes for hpG-CSF
CC   cDNA and genomic clones, both of which are claimed. Specifically claimed
CC   are DNA sequences encoding for (Ala1)hpG-CSF; (Ser 36, 42, 64 and 74)hpG
CC   -CSF and the corresponding Met-1 cpds. The novelty is that hpG-CSF is the
CC   prod. of procaryotic or eucaryotic expression of an exogenous DNA
CC   sequence. The construction of hpG-CSF expression vectors is also
CC   described in the examples. (Updated on 25-MAR-2003 to correct PF field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1415 BP; 265 A; 451 C; 380 G; 319 T; 0 U; 0 Other;

Query Match   99.7%; Score 521.4; DB 1; Length 1415;
Best Local Similarity   99.8%; Pred. No. 5.6e-107;
Matches   522; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy    1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   37 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 96

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGCAGGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97 GTGAGGAAGATCCAGGGCGATGGCGCAGGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 156

Qy  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 216

Qy  181 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGCGGCTGCTTGAGCCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217 CTGAGCAGCTGCCCCAGCCAGGCCCTGCCAGCTGCGGCTGCTTGAGCCAACTCCATAGC 276

Qy  241 GGCCTTTCCTTACCAGGGGTCCTCGGACGCCCTGGAAGGATCCCCGAGGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  277 GGCCTTTCCTTACCAGGGGTCCTCGGACGCCCTGGAAGGATCCCCGAGGTTGGGT 336

Qy  301 CCCACCTTGGACACACTGCAGTGACGTCGCCGACTTTGCCACCACCATCTGGCAGCAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  337 CCCACCTTGGACACACTGCAGTGACGTCGCCGACTTTGCCACCACCATCTGGCAGCAG 396

Qy  361 ATGGAAGAACTGGGAATGGCC... 420
         
Db  397 ATGGAAGAACTGGGAATGGCC... 456

Qy  421 GCCTCTGCTTTCAGCGCCGGCAGGAGGGGGTCCTCCATCTGCAGAGCTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  457 GCCTCTGCTTTCAGCGCCGGCAGGAGGGGGTCCTCCATCTGCAGAGCTTC 516
```

```
Qy    481  CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
           ||||||||||||||||||||||||||||||||||||||||||||
Db    517  CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 559
```

RESULT 12

```
ID   ABX63825   standard; cDNA; 1508 BP.
XX
AC   ABX63825;
XX
DT   26-FEB-2003  (first entry)
XX
DE   Human cDNA #825 differentially expressed in activated vascular tissue.
XX
KW   Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;
KW   hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;
KW   gene therapy; vascular disease; cancer; coronary; artery disease;
KW   hypertension; diabetes; pre-eclampsia; restenosis;
KW   ischaemia-reperfusion injury; stroke.
XX
OS   Homo sapiens.
XX
PN   US2002137081-A1.
XX
PD   26-SEP-2002.
XX
PF   08-JAN-2002; 2002US-00044090.
XX
PR   28-JUL-2000; 2000US-0222469P.
PR   08-JAN-2001; 2001US-0260483P.
XX
PA   (BAND/) BANDMAN O.
XX
PI   Bandman O;
XX
DR   WPI; 2003-110597/10.
XX
PT   Combination for diagnosing, staging, treating, or monitoring the
PT   progression of treatment of a vascular disease, e.g. atherosclerosis,
PT   comprises several cDNAs that are differentially expressed in activated
PT   vascular tissue.
XX
PS   Claim 1; Page; 18pp; English.
XX
CC   This invention relates to a combination comprising several cDNAs that are
CC   differentially expressed in activated vascular tissue. The invention also
CC   discloses a high throughput method for detecting differentially expressed
CC   cDNAs in a sample. The cDNAs of the invention may have
CC   antiarteriosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;
CC   gynaecological; vasotropic and cerebroprotective activities and may be
CC   used in gene therapy. The cDNAs of the invention may be used in a high-
CC   throughput methods for detecting differential expression of one or more
CC   cDNAs in a sample, or screening several molecules or compounds to
CC   identify a molecule or compound that specifically binds a cDNA of the
CC   invention. A protein encoded by the cDNA may be used to screen several
CC   molecules or compounds to identify a ligand that specifically binds to the
CC   protein, or to produce or purify an antibody to the protein that can
CC   be used to detect a protein in a sample or purify a natural or
CC   recombinant protein from a sample. The nucleotides may be useful for
CC   diagnosing, staging, treating, or monitoring the progression of treatment
CC   of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
CC   disease, hypertension, diabetes, pre-eclampsia, ischaemia- reperfusion
CC   injury, restenosis, or stroke. The cDNAs can also be used for large-scale
CC   genetic or gene expression analysis of several new nucleic acid
CC   molecules. Antibodies to the proteins encoded by the cDNAs are useful for
CC   diagnosing pre-pathologic disorders, and chronic or acute diseases
CC   associated with abnormalities in the expression, amount or distribution
CC   of the protein. The present sequence represents a cDNA of the invention
CC   that is differentially expressed in activated vascular tissue. Note: The
CC   sequence data for this patent did not form part of the specification, but
CC   was obtained in electronic format directly from USPTO at
```

```
CC   http.seqdata.uspto.gov/sequence.html?DocID-20020137081
XX
SQ   Sequence 1508 BP; 294 A; 484 C; 402 G; 328 T; 0 U; 0 Other;
```

```
     Query Match           99.7%;  Score 521.4;  DB 7;  Length 1508;
     Best Local Similarity 99.8%;  Pred. No. 5.6e-107;
     Matches 522;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;
```

```
Qy      1  ACCCCCCTGGGCCCTGCCAGCTCCCGCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    114  ACCCCCCTGGGCCCTGCCAGCTCCCGCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 173

Qy     61  GTGAGGAAGATCCAGGGCGATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    174  GTGAGGAAGATCCAGGGCCATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 233

Qy    121  CTGTGCCACCCGGAGGAGCTGGTGCTGCTGCTCTCGGGACACTCTCTGGGCCTCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    234  CTGTGCCACCCGGAGGAGCTGGTGCTGCTGCTCTCGGGACACTCTCTGGGCCTCC 293

Qy    181  CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGTGCAGGCTGCTTGGCAGGCTGGCAACTCCATAGC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    294  CTGAGCAGCTGCCCCAGCCAGGCCCTGCAGTGCAGGCTGCTTGGCAGGCTGGCAACTCCATAGC 353

Qy    241  GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    354  GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 413

Qy    301  CCCACCTTGGACACACTGCAGCTGCACTGCGACGTGCGCCACCATCTGCGACAG 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    414  CCCACCTTGGACACACTGCAGCTGCACTGCGACGTGCGCCACCATCTGCGACAG 473

Qy    361  ATGGAAGAACTGGAAATGGCCCCTGCCCTGCGAGCCCACCCAGGGTGCCCAGGGTGCGCCTTC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    474  ATGGAAGAACTGGAAATGGCCCCTGCCCTGCGAGCCCACCCAGGGTGCCCAGGGTGCGCCTTC 533

Qy    421  GCCTCTGCTTCCAGCGCCGGGCAGGAGGGGATCCCAGTTGCTCCCATCTGCAGAGCTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    534  GCCTCTGCTTCCAGCGCCGGGCAGGAGGGGATCCCAGTTGCTCCCATCTGCAGAGCTTC 593

Qy    481  CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
           |||||||||||||||||||||||||||||||||||||||||||
Db    594  CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 636
```

RESULT 13

```
AAI71848
ID   AAI71848   standard; cDNA; 644 BP.
XX
AC   AAI71848;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Recombinant human granulocyte colony stimulating factor cDNA.
XX
KW   Human; granulocyte colony stimulating factor; rhG-CSF; recombinant; ss.
XX
OS   Homo sapiens.
XX
FH   Key              Location/Qualifiers
FT   CDS              94..621
FT                    /tag= a
FT                    /product= "rhG-CSF"
XX
PN   CN1167150-A.
XX
PD   10-DEC-1997.
XX
PF   05-JUN-1996; 96CN-00106418.
XX
PR   05-JUN-1996; 96CN-00106418.
XX
PA   (JIUY-) JIUYUAN GENE ENG CO LTD HANGZHOU.
```

```
XX
PI   Su Y, Kong T, Wang C;
XX
DR   WPI; 2001-590346/67.
DR   P-PSDB; AAM51536.
XX
PT   Production of recombined human granulocyte colony stimulation factor.
XX
PS   Disclosure; Page 8 (disclosure); 15pp; Chinese.
XX
CC   The invention relates to a method for producing recombinant human
CC   granulocyte colony stimulating factor (rhG-CSF). The method includes the
CC   following steps: (a) reverse transcription-polymerase chain reaction of
CC   the human granulocyte colony stimulating factor gene; (b) transforming
CC   Escherichia coli; (c) renaturation of protein by hollow fibre
CC   ultrafiltration dialysis; (d) passing the renatured protein through ion
CC   exchange chromatography, hydrophobic chromatography and molecular sieve
CC   chromatography which are combined together sequentially; and (e)
CC   purifying so to obtain a high yield of high-purity medicinal rhG-CSF
CC   protein. The present sequence encodes rhG-CSF
XX
SQ   Sequence 644 BP; 128 A; 216 C; 172 G; 128 T; 0 U; 0 Other;

Query Match            98.2%;  Score 513.4;  DB 5;  Length 644;
Best Local Similarity  98.9%;  Pred. No. 3.1e-105;
Matches  517;  Conservative   0;  Mismatches   6;  Indels   0;  Gaps   0;

Qy     1 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    97 ACACCATTAGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 156

Qy    61 GTGAGGAAGATCCAGGGCCGATGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   157 GTGAGGAAGATCCAGGGCCGATGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 216

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   217 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 276

Qy   181 CTGAGCAGCTGCGCCCAGCCAGGCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   277 CTGAGCAGCTGCGCCCAGCCAGGCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 336

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGCGAGGGATCTCCCCGAGTTGGGT 300
           | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   337 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 396

Qy   301 CCCACCTTGGACACACTGCAGCTCGAGCGTCGCCGACTTGCCACCACCATCTGGCAGCAG 360
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   397 CCCACCTTGGACACACTGCAGCTCGAGCGTCGCCGACTTGCCACCACCATCTGGCAGCAG 456

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCAGCCACCCAGGGTGCCATGCCGGCCTTC 420
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   457 ATGGAAGAACTGGGAATGGCCCCTGCCCAGCCACCCAGGGTGCCATGCCGGCCTTC 516

Qy   421 GCCTCTGCTTTCCAGCGCGGGCCGCGGGCAGGAGGGGTCCCAGTTGCCTCCCATCTGCAGAGCTTC 480
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   517 GCCTCTGCTTTCCAGCGCGGGCCGCGGGCAGGAGGGGTCCTGGTTGCCTCCCATCTCCAGAGCTTC 576

Qy   481 CTGGAGGTGTGTACCGGCTTCTACGCCACTTGCCCAGCCCT 523
           | | | | | | | | | | | | | | | | | | | | |
Db   577 CTGGAGGTGTCGTACCGGCTTCTACGCCACTTGCCCAGCCCT 619
```

RESULT 14
AA221169
ID   AA221169 standard; cDNA; 540 BP.
XX
AC   AA221169;
XX
DT   19-NOV-1999 (first entry)
XX
DE   Recombinant human granulocyte colony stimulating factor encoding cDNA.

```
XX
KW   Human; granulocyte colony stimulating factor; G-CSF; rhG-CSF; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..540
FT                  /*tag= a
FT                  /product= "recombinant human G-CSF"
FT                  /transl_except= (pos:13..15,aa:Xaa)
FT                  /note="Xaa represents nothing i.e. nothing is given in
FT                  the protein at this position"
FT   misc_difference 1..12
FT                  /*tag= b
FT                  /note= "this part of the sequence encodes a 4 amino acid
FT                  peptide not present in the wild type protein"
FT   misc_difference 64..66
FT                  /*tag= c
FT                  /note= "alternatively TGC at this position is replaced by
FT                  GCC causing an Ala residue to replace Cys in the protein"
XX
PN   CN1217342-A.
XX
PD   26-MAY-1999.
XX
PF   19-NOV-1997;  97CN-00123162.
XX
PR   19-NOV-1997;  97CN-00123162.
XX
PA   (JIAN/) JIANG Y.
XX
PI   Jiang Y, Dai W;
XX
DR   WPI; 1999-459246/39.
DR   P-PSDB; AAY29929.
XX
PT   Colony stimulating factor for recombination of human granulocytes -
PT   obtained by adding 3-6 amino acid residues at N end of its sequence, in
PT   which at least one amino acid residue is arginine.
XX
PS   Claim 1; Fig 1; 16pp; Chinese.
XX
CC   The present sequence encodes a recombinant human granulocyte colony
CC   stimulating factor (rhG-CSF). The present invention describes rhG-CSF
CC   proteins which are obtained by adding 3-6 amino acid residues at the N-
CC   terminal of the hG-CSF sequence, in which at least one amino acid residue
CC   is arginine. Preferably adding 4-6 amino acid residues including the
CC   sequence of methionine, arginine, glutamic acid and serine at the N-
CC   terminal. Also described are the related sequence, expression carrier,
CC   host cell and medicinal composition. The activity of the recombinant
CC   protein is increased and its stability is increased
XX
SQ   Sequence 540 BP; 90 A; 191 C; 158 G; 101 T; 0 U; 0 Other;

Query Match            97.6%;  Score 510.2;  DB 2;  Length 540;
Best Local Similarity  98.5%;  Pred. No. 1.6e-104;
Matches  515;  Conservative   0;  Mismatches   8;  Indels   0;  Gaps   0;

Qy     1 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    16 ACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 75

Qy    61 GTGAGGAAGATCCAGGGCCGATGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    76 GTGAGGAAGATCCAGGGCCGATGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 135

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   136 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 195

Qy   181 CTGAGCAGCTGCGCCCAGCCAGGCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
           | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
SQ   Sequence 556 BP; 94 A; 197 C; 159 G; 106 T; 0 U; 0 Other;

     Query Match             97.6%;  Score 510.2;  DB 1;  Length 556;
     Best Local Similarity   98.5%;  Pred. No. 1.6e-104;
     Matches  515;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1 ACCCCCTGGGCCTGCCAGCTCCCTGCTCCCCGAGCTTCCTGCTCAAGTGCTTAGAGCAA   60
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     17 ACCCCCTGGGCCTGCCAGCTCCCTGCTCCCCGAGCTTCCTGCTCAAGTGCTTAGAGCAA   76

Qy     61 GTGAGGAAGATCAGGGGCGGATGGCGCAGCGGGCGGCTCCAGGAGAAGCTGTGTGCCACTACAAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 GTGAGGAAGATCAGGGGCGGATGGCGCAGCGGGCGGCTCCAGGAGAAGCTGTGTGCCACTACAAG  136

Qy    121 CTGTGCCACCCCGAGGAGCTGGTCGTCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 CTGTGCCACCCCGAGGAGCTGGTCGTCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC  196

Qy    181 CTGAGCAGCTGCCCAGCCAGGCCCTGCAGCGTGGCAGCGCCCTGGAAGGGATCTCCCCGAGTTGGGT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 CTGAGCTCCTGCCCAGCCAGGCCCTGCAGCGTGGCAGCGCCCTGGAAGGGATATCCCCGAGTTGGGT  256

Qy    241 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGGCCCTGGAAGGGATCTCCCCGAGTTGGGT  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGGCCCTGGAAGGGATATCCCCGAGTTGGGT  316

Qy    301 CCCACCTTGGACACACTGACGTGCGCCGACTTTGCCACCACCATCTGCAGCAG  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 CCCACCTTGGACACACTGACGTGCGACGTTTGCCACCACCATCTGCAGCAG  376

Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCGCAGCCACCAGGGTGCCATGCCGGCCTTC  420
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    377 ATGGAAGAACTGGGAATGGCCCCTGCCGCAGCCACCAGGGTGCCATGCCGGCCTTC  436

Qy    421 GCCTCTGCTTTCCAGCGCGCGGGCAGGAGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC  480
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    437 GCCTCTGCTTTCCAGCGCGCGGGCAGGAGGGTCCTGGTTGCCATCTGCAGAGCTTC  496

Qy    481 CTGGAGGTGCGTACCGCGGTTCTACGCGCCACCTTGCCCAGCCCT  523
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    497 CTGGAGGTGTCGTACCGCGGTTCTACGCGCCACCTTGCCCAGCCCT  539


Search completed: October  9, 2004, 02:07:16
Job time : 402.272 secs
```

```
Db    196 CTGAGCTCCTGCCCAGCCAGGCCCTGCAGCGTGGCAGCGCCCTGGAAGGGATCTCCCCGAGTTGGGT  255

Qy    241 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGGCCCTGGAAGGGATCTCCCCGAGTTGGGT  300
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    256 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGGCCCTGGAAGGGATATCCCCGAGTTGGGT  315

Qy    301 CCCACCTTGGACACACTGACGTGCGCCGACTTTGCCACCACCATCTGGCAGCAG  360
          ||||||||||||||||||||||||| |||||||||||||||||| |||||||||
Db    316 CCCACCTTGGACACACTGACGTGCGACGTTTGCCACCACCATCTGCAGCAG  375

Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCGCAGCCACCAGGGTGCCATGCCGGCCTTC  420
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    376 ATGGAAGAACTGGGAATGGCCCCTGCCGCAGCCACCAGGGTGCCATGCCGGCCTTC  435

Qy    421 GCCTCTGCTTTCCAGCGCGCGGGCAGGAGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC  480
          ||||||||||||||||||||||||||||||||||||| |||||||| ||||||||||||||
Db    436 GCCTCTGCTTTCCAGCGCGCGGGCAGGAGGGTCCTGGTTGCCATCTGCAGAGCTTC  495

Qy    481 CTGGAGGTGCGTACCGCGGTTCTACGCGCCACCTTGCCCAGCCCT  523
          |||||||||| |||||||||||||||||||||||||||||||||||
Db    496 CTGGAGGTGTCGTACCGCGGTTCTACGCGCCACCTTGCCCAGCCCT  538
```

```
RESULT 15
AAN90533
ID   AAN90533 standard; DNA; 556 BP.
XX
AC   AAN90533;
XX
DT   25-MAR-2003  (revised)
DT   27-NOV-1989  (first entry)
XX
DE   Synthetic DNA encoding human granulocyte colony stimulating factor.
XX
KW   Synthetic DNA; human granulocyte colony stimulating factor;
KW   restriction sites; bone marrow precursor cells.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   CDS            14..538
FT                  /*tag= a
XX
PN   GB2213821-A.
XX
PD   23-AUG-1989.
XX
PF   23-DEC-1987;  87GB-00030055.
XX
PR   23-DEC-1987;  87GB-00030055.
XX
PA   (BRBI-) BRIT BIO-TECHN LTD.
XX
PI   Edwards RM;
XX
DR   WPI; 1989-243775/34.
DR   P-PSDB; AAP91070.
XX
PT   Synthetic granulocyte colony stimulating factor gene - incorporating
PT   restriction sites for facilitating cassette mutagenesis and expression.
XX
PS   Claim 2; Page 14; 24pp; English.
XX
CC   The DNA encodes granulocyte colony stimulating factor (G-CSF), with
CC   restriction sites for HindIII, BspMI, ApaI, BstXI, Eco47III, SacI, EcoRV,
CC   AatII, PfiMI, NheI, FspI, BamHI, and EcoRI. G-CSF stimulates
CC   proliferation of specific bone marrow precursor cells and their
CC   differentiation into granulocytes. The gene incorporates useful
CC   restriction sites at frequent intervals to facilitate the cassette
CC   mutagenesis of selected regions. See also AAP91070. (Updated on 25-MAR-
CC   2003 to correct PA field.)
XX
```

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: October 8, 2004, 21:23:54 ; Search time 2723.62 Seconds
                                      (without alignments)
                                      8322.904 Million cell updates/sec

Title: US-10-009-792C-18_COPY_88_610
Perfect score: 523
Sequence: 1 accccccctgggccctgccag........acgccaccttgcccagccct 523

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151699 5 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
      1: gb_ba:*
      2: gb_htg:*
      3: gb_in:*
      4: gb_om:*
      5: gb_ov:*
      6: gb_pat:*
      7: gb_ph:*
      8: gb_pl:*
      9: gb_pr:*
     10: gb_ro:*
     11: gb_sts:*
     12: gb_sy:*
     13: gb_un:*
     14: gb_vi:*
     15: gb_ba:*
     16: gb_fun:*
     17: gb_hum:*
     18: gb_in:*
     19: em_mu:*
     20: em_om:*
     21: em_ov:*
     22: em_ov:*
     23: em_pat:*
     24: em_ph:*
     25: em_pl:*
     26: em_ro:*
     27: em_sts:*
     28: em_un:*
     29: em_vi:*
     30: em_htg_hum:*
     31: em_htg_inv:*
     32: em_htg_other:*
     33: em_htg_mus:*
     34: em_htg_pln:*
     35: em_htg_rod:*
     36: em_htg_vrt:*
     37: em_htg_vrt:*
     38: em_sy:*
     39: em_htgo_hum:*
     40: em_htgo_mus:*
     41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	523	100.0	525	6	AR024358	AR024358 Sequence
2	523	100.0	525	6	AR091731	AR091731 Sequence
3	523	100.0	525	6	E01731	E01731 cDNA encodi
4	523	100.0	525	6	E02573	E02573 DNA encodin
5	523	100.0	525	6	E07164	E07164 cDNA encodi
6	523	100.0	525	6	E09431	E09431 cDNA encodi
7	523	100.0	525	6	E15131	E15131 Human mRNA
8	523	100.0	525	6	I71150	I71150 Sequence 1
9	523	100.0	525	6	I83709	I83709 Sequence 1
10	523	100.0	525	6	AR363055	AR363055 Sequence
11	523	100.0	660	9	HUMGCSF	M17706 Human granu
12	523	100.0	1498	9	HSGCSFR1	X03655 Human mRNA
13	523	100.0	1521	6	E01219	E01219 cDNA encodi
14	523	100.0	1525	6	I05333	I05333 Sequence 11
15	523	100.0	1525	6	I08669	I08669 Sequence 1
16	523	100.0	1525	6	I09205	I09205 Sequence 5
17	521.4	99.7	1414	9	HUMGCSFB	M13008 Human granu
18	521.4	99.7	1671	9	BC033245	BC033245 Homo sapi
19	520.4	99.5	522	6	E01631	E01631 DNA encodin
20	520.4	99.5	522	6	E08529	E08529 DNA encodin
21	519.8	99.4	1415	6	E37811	E37811 Vector. 6/2
22	519.8	99.4	1415	6	I01623	I01623 Sequence 1
23	519.8	99.4	1415	6	BD000076	BD000076 Cloning m
24	519.8	99.4	1415	6	BD178096	BD178096 Cloning m
25	510.2	97.6	541	6	A76866	A76866 Sequence 4
26	510.2	97.6	777	6	AX027608	AX027608 Sequence
27	510.2	97.6	2382	6	A76771	A76771 Sequence 1
28	510.2	97.6	2382	6	I64766	I64766 Sequence 1
29	509.2	97.4	2455	6	A76772	A76772 Sequence 2
30	509.2	97.4	2455	6	A76867	A76867 Sequence 5
31	509.2	97.4	2455	6	I64767	I64767 Sequence 3
32	507	96.9	525	6	AR123945	AR123945 Sequence
33	506	96.7	522	6	AX398641	AX398641 Sequence
34	506	96.7	522	6	AX398644	AX398644 Sequence
35	504.4	96.4	522	6	AX398637	AX398637 Sequence
36	504.4	96.4	522	6	AX398638	AX398638 Sequence
37	504.4	96.4	522	6	AX398639	AX398639 Sequence
38	504.2	96.4	519	6	AX027604	AX027604 Sequence
39	504.2	96.4	522	6	AX027606	AX027606 Sequence
40	503.8	96.3	546	6	AR202281	AR202281 Sequence
41	503.8	96.3	546	6	AR223283	AR223283 Sequence
42	503.8	96.3	921	6	AR202223	AR202223 Sequence
43	503.8	96.3	921	6	AR202226	AR202226 Sequence
44	503.8	96.3	921	6	AR202229	AR202229 Sequence
45	503.8	96.3	921	6	AR223225	AR223225 Sequence

ALIGNMENTS

RESULT 1
AR024358

LOCUS   AR024358   525 bp   DNA   linear   PAT 05-DEC-1998
DEFINITION Sequence 1 from patent US 5795968.
ACCESSION   AR024358
VERSION   AR024358.1   GI:3977652
KEYWORDS
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 525)
  AUTHORS   Kuga.T., Miyaji.H., Sato,M., Okabe,M., Morimoto,M., Itoh.S.,
            Yamasaki,M., Yokoo,Y., Yamaguchi,K., Yoshida,H. and Komatsu,Y.
  TITLE   Polypeptide derivatives of human granilocyte colony stimulating
            factor

JOURNAL   Patent: US 5795968-A 1 18-AUG-1998;
FEATURES   Location/Qualifiers
     source   1..525
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   100.0%; Score 523; DB 6; Length 525;
Best Local Similarity   100.0%; Pred. No. 2.5e-86;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy    61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCGTTCCAGGAGAGCTGTGTGCCACTACAAG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCGTTCCAGGAGAGCTGTGTGCCACTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGAGCTGGCTGCTCGGACACTCTCTGGGCATCCCCTGGCTCCC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGAGCTGGCTGCTCGGACACTCTCTGGGCATCCCCTGGCTCCC 180

Qy   181 CTGAGCAGTGCCCCAGCCAGGCCCTGCTTGAGCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGTGCCCCAGCCAGGCCCTGCTTGAGCAACTCCATAGC 240

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300

Qy   301 CCCACCTTGGACACACTGCAGCTGCGCCGACGTTGCCACCATCTGCGACAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACACTGCAGCTGCGCCGACGTTGCCACCATCTGCGACAG 360

Qy   361 ATGAAGAAGACTGGGAATGGCCCTGCAGCCCTGCAGGGTGCCATGCCGCCTTC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGAAGAAGACTGGGAATGGCCCTGCAGCCCTGCAGGGTGCCATGCCGCCTTC 420

Qy   421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGGTCCTCCAGTTGCCTCCCATCTGCGAGAGCTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGGTCCTCCAGTTGCCTCCCATCTGCGAGAGCTTC 480

Qy   481 CTGAGGGTGTCGTACCGGCGTTCTACGCCCACCTTGCCCAGCCCT 523
         |||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGAGGGTGTCGTACCGGCGTTCTACGCCCACCTTGCCCAGCCCT 523
```

RESULT 2
AR091731
LOCUS   AR091731   525 bp   DNA   linear   PAT 07-SEP-2000
DEFINITION   Sequence 1 from patent US 5994518.
ACCESSION   AR091731
VERSION   AR091731.1 GI:10018485
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 525)
  AUTHORS   Kuga,T., Miyaji,H., Sato,M., Okabe,M., Morimoto,M., Itoh,S.,
             Yamasaki,M., Yokoo,Y., Yamaguchi,K., Yoshida,H. and Komatsu,Y.
  TITLE   Method of producing a polypeptide having human granulocyte colony
             stimulating factor activity
  JOURNAL   Patent: US 5994518-A 1 30-NOV-1999;
FEATURES   Location/Qualifiers
     source   1..525
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   100.0%; Score 523; DB 6; Length 525;
Best Local Similarity   100.0%; Pred. No. 2.5e-86;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy    61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCGTTCCAGGAGAGCTGTGTGCCACTACAAG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GTGAGGAAGATCCAGGGCGATGGCCAGCGCGTTCCAGGAGAGCTGTGTGCCACTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGAGCTGGCTGCTCGGACACTCTCTGGGCATCCCCTGGCTCCC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGAGCTGGCTGCTCGGACACTCTCTGGGCATCCCCTGGCTCCC 180

Qy   181 CTGAGCAGTGCCCCAGCCAGGCCCTGCTTGAGCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGTGCCCCAGCCAGGCCCTGCTTGAGCAACTCCATAGC 240

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300

Qy   301 CCCACCTTGGACACACTGCAGCTGCGCCGACGTTGCCACCATCTGCGACAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACACTGCAGCTGCGCCGACGTTGCCACCATCTGCGACAG 360

Qy   361 ATGAAGAAGACTGGGAATGGCCCTGCAGCCCTGCAGGGTGCCATGCCGCCTTC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGAAGAAGACTGGGAATGGCCCTGCAGCCCTGCAGGGTGCCATGCCGCCTTC 420

Qy   421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGGTCCTCCAGTTGCCTCCCATCTGCGAGAGCTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGGTCCTCCAGTTGCCTCCCATCTGCGAGAGCTTC 480

Qy   481 CTGAGGGTGTCGTACCGGCGTTCTACGCCCACCTTGCCCAGCCCT 523
         |||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGAGGGTGTCGTACCGGCGTTCTACGCCCACCTTGCCCAGCCCT 523
```

RESULT 3
E01731
LOCUS   E01731   525 bp   RNA   linear   PAT 29-SEP-1997
DEFINITION   cDNA encoding hG-CSF form human periferal macrophage.
ACCESSION   E01731
VERSION   E01731.1 GI:2169984
KEYWORDS   JP 198267292-A/1.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 525)
  AUTHORS   Kuga,T., Komatsu,Y., Miyaji,H., Sato,M., Okabe,M., Morimoto,M.,
             Itou,S., Yamazaki,M., Yokoo,Y. and Yamaguchi,K.
  TITLE   NOVEL POLYPEPTIDE
  JOURNAL   Patent: JP.198267292-A 1 04-NOV-1988;
             KYOWA HAKKO KOGYO CO LTD
COMMENT   OS   Human
             PN   JP 198267292-A/1
             PD   04-NOV-1988
             PF   23-DEC-1987 JP 1987326384
             PR   23-DEC-1986 JP 86P 306799
             PI   KUGA TETSUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI, PI
             OKABE MASAMI, MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIHARU,
             PI   YAMAGUCHI KAZUO
             PC   C12P21/02, C07K13/00, C12N15/00//A61K37/02, A61K37/02,
             PC   A61K37/02,
             PC   (C12P21/02, C07K13/00, C12N15/00//A61K37/02, A61K37/02,
             PC   (C12P21/02, C12R1:91);
             CC   strandedness: Double;
             CC   topology: Linear;
             CC   hypothetical: No;
             CC   *source: cell_line=Periferal macrophage;
             FH   Key   Location/Qualifiers
             FH

```
                    PI   ITOU SEIGA
                    PC   C12N9/72,C07K13/00,C07K15/14,C12N1/21,C12N5/10,C12N9/64, PC
                    C12N15/27,
                    PC   C12N15/58,C12N15/70,C12N15/85,C12P21/02,C12P21/02,(C12N1/21,
                    PC   C12R1:19),
                    PC   (C12P21/02,C12R1:19),(C12P21/02,C12R1:91);
                    CC   strandedness: Double;
                    CC   topology: Linear;
                    CC   hypothetical: No;
                    CC   anti-sense: No;
                    CC   *source: cell type=lymphoblast LukII;
                    CC   *source: clone=pCSF1-2;
                    FH        Key            Location/Qualifiers
                    FH
FT   mat_peptide    FT   mat_peptide     1..522
Location/Qualifiers FT                   /product='human colony-stimulating  FT
     1..525                                  factor(hG-CSF)'.
     /product='mature peptide of hG-CSF'.    Location/Qualifiers
     /organism="Homo sapiens"       source   1..525
     /mol_type="genomic RNA"                 /organism="Homo sapiens"
     /db_xref="taxon:9606"                   /mol_type="genomic RNA"
                                             /db_xref="taxon:9606"
ORIGIN                              ORIGIN

Query Match  100.0%;  Score 523;  DB 6;  Length 525;      Query Match  100.0%;  Score 523;  DB 6;  Length 525;
Best Local Similarity 100.0%; Pred. No. 2.5e-86;          Best Local Similarity 100.0%; Pred. No. 2.5e-86;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;   Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 ACCCCCCTGGGCCCTGCCAGCTCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60     Qy  1 ACCCCCCTGGGCCCTGCCAGCTCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1 ACCCCCCTGGGCCCTGCCAGCTCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60     Db  1 ACCCCCCTGGGCCCTGCCAGCTCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy 61 GTGAGGAAGATCCAGGGCGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120   Qy 61 GTGAGGAAGATCCAGGGCGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GTGAGGAAGATCCAGGGCGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120   Db 61 GTGAGGAAGATCCAGGGCGCGATGGCGCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy 121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180   Qy 121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180   Db 121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGGCTCCC 180

Qy 181 CTGAGCAGCTGCCCAGCCAGGCCCTGCCAGCTGCGTTGAGCAACTCCATAGC 240   Qy 181 CTGAGCAGCTGCCCAGCCAGGCCCTGCCAGCTGCGTTGAGCAACTCCATAGC 240
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CTGAGCAGCTGCCCAGCCAGGCCCTGCCAGCTGCGTTGAGCAACTCCATAGC 240   Db 181 CTGAGCAGCTGCCCAGCCAGGCCCTGCCAGCTGCGTTGAGCAACTCCATAGC 240

Qy 241 GGCCTTTTCCTCTACCAGGGGCTCGCAGCGGGGGATCTCCCCGAGTTGGGT 300   Qy 241 GGCCTTTTCCTCTACCAGGGGCTCGCAGCGGGGGATCTCCCCGAGTTGGGT 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GGCCTTTTCCTCTACCAGGGGCTCGCAGCGGGGGATCTCCCCGAGTTGGGT 300   Db 241 GGCCTTTTCCTCTACCAGGGGCTCGCAGCGGGGGATCTCCCCGAGTTGGGT 300

Qy 301 CCCACCTTGGACACACTGGACGTGCGCGGACTTTGCCGCCATCTGGCAGCAG 360   Qy 301 CCCACCTTGGACACACTGGACGTGCGCGGACTTTGCCGCCATCTGGCAGCAG 360
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CCCACCTTGGACACACTGGACGTGCGCGGACTTTGCCGCCATCTGGCAGCAG 360   Db 301 CCCACCTTGGACACACTGGACGTGCGCGGACTTTGCCGCCATCTGGCAGCAG 360

Qy 361 ATGGAAGAACTGGGAATGGCCCTGCCCTCCAGGGTGCCATGCCGGCCTTC 420   Qy 361 ATGGAAGAACTGGGAATGGCCCTGCCCTCCAGGGTGCCATGCCGGCCTTC 420
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 ATGGAAGAACTGGGAATGGCCCTGCCCTCCAGGGTGCCATGCCGGCCTTC 420   Db 361 ATGGAAGAACTGGGAATGGCCCTGCCCTCCAGGGTGCCATGCCGGCCTTC 420

Qy 421 GCCTCTGCTTTCAGCGCGCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480   Qy 421 GCCTCTGCTTTCAGCGCGCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GCCTCTGCTTTCAGCGCGCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480   Db 421 GCCTCTGCTTTCAGCGCGCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480

Qy 481 CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCCT 523   Qy 481 CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCCT 523
       ||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||||||||||||||||||||||||||||||||||||||
Db 481 CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCCT 523   Db 481 CTGGAGGTGTCGTACCGCGTTCTACGCGCCACCTTGCCCAGCCCT 523
```

```
RESULT 4
E02573
LOCUS        E02573           525 bp    RNA      linear      PAT 29-SEP-1997
DEFINITION   DNA encoding human colony-stimulating factor(hG-CSF).
ACCESSION    E02573
VERSION      E02573.1   GI:2170803
KEYWORDS     JP 1990227075-A/1.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 525)
  AUTHORS    Sasaki,K., Nishi,T., Yasumura,S., Sato,M. and Itou,S.
  TITLE      NEW POLYPEPTIDE
  JOURNAL    Patent: JP 1990227075-A 1 10-SEP-1990;
             KYOWA HAKKO KOGYO CO LTD
COMMENT      OS   Homo sapiens (human)
             PN   JP 1990227075-A/1
             PD   10-SEP-1990
             PF   28-SEP-1989 JP 1989253097
             PR   29-SEP-1988 JP 88P  245705
             PI   SASAKI KATSUTOSHI, NISHI TATSUYA, YASUMURA SHIGEYOSHI, PI
             SATO MORIYUKI,
```

```
RESULT 5
E07164
LOCUS        E07164           525 bp    RNA      linear      PAT 29-SEP-1997
DEFINITION   cDNA encoding human G-CSF.
ACCESSION    E07164
VERSION      E07164.1   GI:2175311
KEYWORDS     JP 1994092994-A/1.
SOURCE       Homo sapiens (human)
```

```
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 525)
AUTHORS    Kuga,T., Komatsu,Y., Miyaji,H., Sato,M., Okabe,M., Morimoto,M.,
           Itou,S., Yamazaki,M., Yokoo,Y. and Yamaguchi,K.
TITLE      NEW POLYPEPTIDE
JOURNAL    Patent: JP 1994092994-A 1 05-APR-1994;
           KYOWA HAKKO KOGYO CO LTD
COMMENT    OS   Homo sapiens (human)
           PN   JP 1994092994-A/1
           PD   05-APR-1994
           PF   23-DEC-1987 JP 1992214376
           PR   23-DEC-1986 JP 86P   306799
           PI   KUGA TETSUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI, PI
           OKABE MASAMI,
           PI   MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIHARU,
           PI   YAMAGUCHI KAZUO
           PC   C07K13/00,A61K37/02,C12N1/21,C12N15/27,C12P21/02,(C12N1/21, PC
           C12R1:19);
           PC   (C12P21/02,C12R1:19);
           CC   strandedness: Double;
           CC   topology: Linear;
           CC   hypothetical: No;
           CC   anti-sense: No;
           FH   Key             Location/Qualifiers
           FH
           FT   source          1..525
           FT                   /organism='Homo sapiens'
           FT                   /cell_type='macrophage'
           FT                   /clone='pCSF1-2'
           FT   CDS             1..525
           FT                   /product='human G-CSF'.
FEATURES            Location/Qualifiers
     source         1..525
                    /organism="Homo sapiens"
                    /mol_type="genomic RNA"
                    /db_xref="taxon:9606"
ORIGIN

Query Match       100.0%;  Score 523;  DB 6;  Length 525;
Best Local Similarity  100.0%;  Pred. No. 2.5e-86;
Matches  523;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy    61 GTGAGGAAGATCCAGGGCGATGGGCCAGCGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GTGAGGAAGATCCAGGGCGATGGGCCAGCGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGACACTCTCTGGGCATCCCCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGACACTCTCTGGGCATCCCCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCAGCCAGGCCCTGCCCAGCTGCTTGAGCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCAGCCAGGCCCTGCCCAGCTGCTTGAGCAACTCCATAGC 240

Qy   241 GGCCTTTTCCTCTACCAGGGGATCTCCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTTCCTCTACCAGGGGATCTCCAGGCCCTGAAGGGATCTCCCCGAGTTGGGT 300

Qy   301 CCCACCTGGACACACTGGACGTGACGTGACGTGCCGACTTTGCCACCACATCTGGCAGCAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACACTGGACGTGACGTGCCCACCACATCTGGCAGCAG 360

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCAGCCCACCCAGGGTGCGCGGCCTTC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCAGCCCACCCAGGGTGCGCGGCCTTC 420

Qy   421 GCCTCTGCTTCCAGCGCCGGCCAGGAGGGGTCCTAGTTGCTCCCATCTGCAGAGCTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCTTCCAGCGCCGGCCAGGAGGGGTCCTAGTTGCTCCCATCTGCAGAGCTTC
```

```
Db    421 GCCTCTGCTTCCAGCGCCGGCCAGGAGGGGTCCTAGTTGCTCCCATCTGCAGAGCTTC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 CTGGAGGTGTCGTACCGGCGTTCTACGCGCCACCTTGCCCAGCCCT 523
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGGAGGTGTCGTACCGGCGTTCTACGCGCCACCTTGCCCAGCCCT 523

RESULT 6
E09431
LOCUS       E09431                   525 bp    RNA    linear    PAT 29-SEP-1997
DEFINITION  cDNA encoding hG-CSF.
ACCESSION   E09431
VERSION     E09431.1  GI:22026058
KEYWORDS    JP 199149798-A/1.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 525)
AUTHORS     Kuga,T., Komatsu,Y., Miyaji,H., Sato,M., Okabe,M., Morimoto,M.,
            Itou,S., Yamazaki,M., Yokoo,Y. and Yamaguchi,K.
TITLE       NEW POLYPEPTIDE
JOURNAL     Patent: JP 199149798-A 1 13-JUN-1995;
            KYOWA HAKKO KOGYO CO LTD
COMMENT     OS   Homo sapiens (Human)
            PN   JP 199149798-A/1
            PD   13-JUN-1995
            PF   08-AUG-1994 JP 1994185787
            PR   23-DEC-1986 JP 86P   306799
            PI   KUGA TETSUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI, PI
            OKABE MASAMI,
            PI   MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIHARU,
            PI   YAMAGUCHI KAZUO
            PC   C07K14/535,C12N1/21,C12N15/09,C12P21/02//A61K38/00,(C12N1/21,
            PC   C12R1:19),
            PC   (C12P21/02,C12R1:19);
            CC   strandedness: Single;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1..525
            FT                   /organism='Homo sapiens'.
FEATURES            Location/Qualifiers
     source         1..525
                    /organism="Homo sapiens"
                    /mol_type="genomic RNA"
                    /db_xref="taxon:9606"
ORIGIN

Query Match      100.0%;  Score 523;  DB 6;  Length 525;
Best Local Similarity  100.0%;  Pred. No. 2.5e-86;
Matches  523;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGTTCCTGCTCAAGTGCTTAGAGCAA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ACCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy    61 GTGAGGAAGATCCAGGGCGATGGGCCAGCGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GTGAGGAAGATCCAGGGCGATGGGCCAGCGGCCAGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGACACTCTCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGACACTCTCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCTGCCAGGCCCTGCCCAGCTGCTTGAGCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCTGCCAGGCCCTGCCCAGCTGCTTGAGCAACTCCATAGC 240

Qy   241 GGCCTTTCCTTCCTACCAGGGGCTCCTCCAGGGGGATCATCCCTGGGCTCCC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTCCTTCCTACCAGGGGCTCCTCCAGGGGGATCATCCCTGGGCTCCC 300

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGACACTCTCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGACACTCTCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCTGCCAGGCCCTGCCCAGCTGCTTGAGCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCTGCCAGGCCCTGCCCAGCTGCTTGAGCAACTCCATAGC 240

Qy   241 GGCCTTTTCCTCTACCAGGGGCTCCTGGAAGGGATCTCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTTCCTCTACCAGGGGCTCCTGGAAGGGATCTCCCCGAGTTGGGT 300
```

RESULT 7
E15131

```
LOCUS       E15131          525 bp    DNA    linear   PAT 28-JUL-1999
DEFINITION  Human mRNA for granulocyte-colony stimulating factor.
ACCESSION   E15131
VERSION     E15131.1  GI:5709814
KEYWORDS    JP 1998052281-A/1.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 525)
AUTHORS     Kuga,T., Komatsu,Y., Miyaji,H., Sato,M., Okabe,M., Morimoto,M.,
            Itou,S., Yamazaki,M., Yokoo,Y. and Yamaguchi,K.
TITLE       NEW POLYPEPTIDE
JOURNAL     Patent: JP 1998052281-A 1 24-FEB-1998;
            KYOWA HAKKO KOGYO CO LTD
COMMENT     OS   Homo sapiens (human)
            PN   JP 1998052281-A/1
            PD   24-FEB-1998
            PF   23-DEC-1987 JP 1997114630
            PR   23-DEC-1986 JP 86P 306799
            PI   KUGA TEISUO, KOMATSU YUKI, MIYAJI HIROMASA, SATO MORIYUKI, PI
            OKABE MASAMI,
            PI   MORIMOTO MAKOTO, ITOU SEIGA, YAMAZAKI MOTOO, YOKOO YOSHIHARU,
            PI   YAMAGUCHI KAZUO
            PC   C12N15/09,A61K38/00,C07K14/535,C12N1/21,C12P21/02,(C12P21/02,
            PC   C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            FH   Key            Location/Qualifiers
            FH
            FT   source         1..525
            FT                  /organism='Homo sapiens'
            FT                  /cell_type='peripheral blood macropharge' FT
            FT   mat_peptide    1..522
            FT                  /product='G-CSF'.
FEATURES             Location/Qualifiers
     source          1..525
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Query Match   100.0%; Score 523; DB 6; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.5e-86;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  60

Qy   61 GTGAGGAAGATCCAGGGCGCGATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGCGATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGGCAGCACTCCTGGGCATCCCTGGGCTCCC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGGCAGCACTCCTGGGCATCCCTGGGCTCCC 180

Qy  181 CTGAGCAGCTGCCCTGCCAGGCCCTGCAGCTGCAGGCTGCTGGAGCCAACTCCATAGC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGAGCAGCTGCCCTGCCAGGCCCTGCAGCTGCAGGCTGCTGGAGCCAACTCCATAGC 240

Qy  241 GGCCTTTCCTCTACCAGGGGATCCCCGAGTTGGGT 300
        |||||||||||||||||||||||||||||||||||
Db  241 GGCCTTTCCTCTACCAGGGGATCCCCGAGTTGGGT 300

Qy  301 CCCACCTTGGACACACTGCAGCTGGACGTGCCGACTTTGCCACCATCTGGCAGCAG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CCCACCTTGGACACACTGCAGCTGGACGTGCCGACTTTGCCACCATCTGGCAGCAG 360

Qy  361 ATGGAAGAACTGGGAATGGCCCGCCTGCAGCCCCACCCAGGGTGCCATGCCGGCCTTC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 ATGGAAGAACTGGGAATGGCCCGCCTGCAGCCCCACCCAGGGTGCCATGCCGGCCTTC 420

Qy  421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GCCTCTGCTTTCCAGCGCCGGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCGAGAGCTTC 480

Qy  481 CTGGAGGTGTCGTACCGGGTTCTACGCCACCTTGCCCAGCCCT 523
        |||||||||||||||||||||||||||||||||||||||||||
Db  481 CTGGAGGTGTCGTACCGGGTTCTACGCCACCTTGCCCAGCCCT 523
```

RESULT 8
I71150

```
LOCUS       I71150          525 bp    DNA    linear   PAT 03-APR-1998
DEFINITION  Sequence 1 from patent US 5681720.
ACCESSION   I71150
VERSION     I71150.1  GI:3007285
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 525)
AUTHORS     Kuga,T., Miyaji,H., Sato,M., Okabe,M., Morimoto,M., Itoh,S.,
            Yamasaki,M., Yokoo,Y., Yamaguchi,K., Yoshida,H. and Komatsu,Y.
TITLE       DNA encoding human granulocyte colony stimulating factor plasmids
            and host cells comprising same, and methods of expressing the
            encoded polypeptide
JOURNAL     Patent: US 5681720-A 1 28-OCT-1997;
FEATURES             Location/Qualifiers
     source          1..525
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Query Match   100.0%; Score 523; DB 6; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.5e-86;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCCCCTGGGCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA  60

Qy   61 GTGAGGAAGATCCAGGGCGCGATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGCGATGGCGCAGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120

Qy  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGGCAGCACTCCTGGGCATCCCTGGGCTCCC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTGGGCAGCACTCCTGGGCATCCCTGGGCTCCC 180

Qy  181 CTGAGCAGCTGCCCTGCCAGGCCCTGCAGCTGCAGGCTGCTGGAGCCAACTCCATAGC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGAGCAGCTGCCCTGCCAGGCCCTGCAGCTGCAGGCTGCTGGAGCCAACTCCATAGC 240
```

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGCTGGAGGGATCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTCCTCTACCAGGGGCTCCTGCTGGAGGGATCCCCGAGTTGGGT 300

Qy   301 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTTGCCACCACATCTGGCAGCAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACACTGCAGCTGGACGTCGCCGACTTTGCCACCACATCTGGCAGCAG 360

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCTGCAGCCCAGGGTGCCATGCCGGCCTTC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGGAAGAACTGGGAATGGCCCCTGCCTGCAGCCCAGGGTGCCATGCCGGCCTTC 420

Qy   421 GCCTCTGCTTTCCAGCGCCGGGCCAGGAGGGGTCCTAGTTGCCTCCATCTGCAGAGCTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCTTTCCAGCGCCGGGCCAGGAGGGGTCCTAGTTGCCTCCATCTGCAGAGCTTC 480

Qy   481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523

RESULT 9
I83709
LOCUS   I83709   525 bp   DNA   linear   PAT 10-AUG-1998
DEFINITION Sequence 1 from patent US 5714581.
ACCESSION   I83709
VERSION   I83709.1 GI:3407239
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 525)
AUTHORS   Kuga,T., Miyaji,H., Sato,M., Okabe,M., Morimoto,M., Itoh,S.,
            Yamasaki,M., Yokoo,Y., Yamaguchi,K., Yoshida,H. and Komatsu,Y.
TITLE   Polypeptide derivatives of human granulocyte colony stimulating
            factor
JOURNAL   Patent: US 5714581-A 1 03-FEB-1998;
FEATURES   Location/Qualifiers
source   1..525
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN

Query Match   100.0%; Score 523; DB 6; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.5e-86;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACCCCCTGGGCCTGCCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCTGGGCCTGCCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCCAGCGCTCCTCCAGGAGAGCGTGTGTCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGGCCAGCGCTCCTCCAGGAGAGCGTGTGTCCACCTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGAGCTGGTGTCGTCCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGAGCTGGTGTCGTCCGGACACTCTCTGGGCATCCCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCTGGCAGGCCCTGCTGCTTGAGCCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCCAGCCAGGCTGGCAGGCCCTGCTGCTTGAGCCAACTCCATAGC 240

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGCTGGAGGGATCTCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTCCTCTACCAGGGGCTCCTGCTGGAGGGATCTCCCCGAGTTGGGT 300

Qy   301 CCCACCTTGGACACTGAGCTGGACGTGCCGCGCGACTTTGCCACCACATCTGGCAGCAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACTGAGCTGGACGTGCCGCGCGACTTTGCCACCACATCTGGCAGCAG 360

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCTGCAGCCCAGGGTGCCATGCCGGCCTTC 420
         ||||||||||||||||||||||||||||||

Db   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTC 420

Qy   421 GCCTCTGCTTTCCAGCGTCCAGCCCTTTCACGCCACCTTGCCCCAGCCCT 523
         |||||||||||||||||||||||
Db   421 GCCTCTGCTTTCCAGCGTCCAGCCCTTTCACGCCACCTTGCCCCAGCCCT 523

Qy   481 CTGGAGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT 523
         ||||||||||||||||||||||||||||
Db   481 CTGGAGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT 523

RESULT 10
AR363055
LOCUS   AR363055   525 bp   DNA   linear   PAT 03-SEP-2003
DEFINITION Sequence 25 from patent US 5194592.
ACCESSION   AR363055
VERSION   AR363055.1 GI:34423821
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 525)
AUTHORS   Yoshida,H.
TITLE   Monoclonal antibodies to novel polypeptide derivatives of human
            granulocyte colony stimulating factor
JOURNAL   Patent: US 5194592-A 25 16-MAR-1993;
FEATURES   Location/Qualifiers
source   1..525
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN

Query Match   100.0%; Score 523; DB 6; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.5e-86;
Matches 523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ACCCCCTGGGCCTGCCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCCCCTGGGCCTGCCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60

Qy   61 GTGAGGAAGATCCAGGGCGATGGGCCAGCGCTCCAGGAGAGCGTGTGCCACCTACAAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGAGGAAGATCCAGGGCGATGGGCCAGCGCTCCAGGAGAGCGTGTGCCACCTACAAG 120

Qy   121 CTGTGCCACCCCGAGGAGAGCTGGTGTCGTCCGGACACTCTCTGGGCATCCCTGGGCTCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGTGCCACCCCGAGGAGAGCTGGTGTCGTCCGGACACTCTCTGGGCATCCCTGGGCTCCC 180

Qy   181 CTGAGCAGCTGCCCCAGCCAGGCTGGCAGGCCCTGCTGCTTGAGCCAACTCCATAGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGAGCAGCTGCCCCAGCCAGGCTGGCAGGCCCTGCTGCTTGAGCCAACTCCATAGC 240

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGCTGGAAGGGATCTCCCCGAGTTGGGT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGCCTTTCCTCTACCAGGGGCTCCTGCTGGAAGGGATCTCCCCGAGTTGGGT 300

Qy   301 CCCACCTTGGACACTGCGAGCTGCCGACTTGCCGACTTTGCCACCATCTGGCAGCAG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCACCTTGGACACTGCGAGCTGCCGACTTGCCGACTTTGCCACCATCTGGCAGCAG 360

Qy   361 ATGGAAGAACTGGGAATGGCCCATGCCCTGCAGCCCCAGGGTGCCATGCCGGCCTTC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGGAAGAACTGGGAATGGCCCATGCCCTGCAGCCCCAGGGTGCCATGCCGGCCTTC 420

Qy   421 GCCTCTGCTTTCCAGCGCCCGGGCCAGGAGGGGTCCTCCTAGTTGCCTCCATCTGCAGAGCTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCTCTGCTTTCCAGCGCCCGGGCCAGGAGGGGTCCTCCTAGTTGCCTCCATCTGCAGAGCTTC 480

Qy   481 CTGGAGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT 523
         ||||||||||||||||||||||||||||
Db   481 CTGGAGTGTCGTACCGGTTCTACGCCACCTTGCCCAGCCCT 523

RESULT 11
HUMGCSF
LOCUS   HUMGCSF   660 bp   mRNA   linear   PRI 27-APR-1993
DEFINITION   Human granulocyte colony stimulating factor mRNA, complete cds.
ACCESSION   M17706
VERSION   M17706.1 GI:183040
KEYWORDS   granulocyte colony stimulating factor.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 660)
AUTHORS   Devlin,J.J., Devlin,P.E., Myambo,K., Lilly,M.B., Rado,T.A. and
               Warren,M.K.
TITLE   Expression of granulocyte colony-stimulating factor by human cell
               lines
JOURNAL   J. Leukoc. Biol. 41 (4), 302-306 (1987)
MEDLINE   87196936
PUBMED   3494801
COMMENT   Original source text: Human,MIA PaCa-2 cell line, cDNA to mRNA,
               (library of Kawasaki et al.), clone pP12.
FEATURES   Location/Qualifiers
     source   1..660
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS   43..657
                     /note="granulocyte colony stimulating factor precursor"
                     /codon_start=1
                     /protein_id="AAA35882.1"
                     /db_xref="GI:183041"
                     /translation="MAGPATQSPMKLMALQLLLWHSALWTVQEATPLGPASSLPQSFL
                     LKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLA
                     GCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPAL
                     QPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP"
     sig_peptide   43..132
                     /note="granulocyte colony stimulating factor signal
                     peptide"
     mat_peptide   133..654
                     /note="granulocyte colony-stimulating factor"
ORIGIN   Unreported.

Query Match   100.0%; Score 523; DB 9; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.4e-86;
Matches   523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACCCCCCTGGGCCCTGCAGCTCCTGCCCAGAGCTCCTCTGCTCAAGTGCTTAGAGCAA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133 ACCCCCCTGGGCCCTGCAGCTCCTGCCCAGAGCTCCTCTGCTCAAGTGCTTAGAGCAA 192

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGGCTCCAGGAGAAGTCGTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193 GTGAGGAAGATCCAGGGCGATGGCGCAGCGGCTCCAGGAGAAGTCGTGTGCCACCTACAAG 252

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTTCTGGGCTCCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTTCTGGGCTCCC 312

Qy   181 CTGAGGAGCAGTGCAATGGCCCTGCCCCAGGGCCCTGCTTGAGGCCAACTCCATAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313 CTGAGGAGCAGTGCAATGGCCCTGCCCCAGGGCCCTGCTTGAGGCCAACTCCATAGC 372

Qy   241 GGCCTTTCCTCTACCAGGGGCTCCTGAGGGCCTTGGAAGGGGATCTCCCCGAGTTGGGT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   373 GGCCTTTCCTCTACCAGGGGCTCCTGAGGGCCTTGGAAGGGGATCTCCCCGAGTTGGGT 432

Qy   301 CCCACCTTGACACACTGCAGTGACGTCGCCGACTTTGCCCACCACCATCTGGCAGCAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433 CCCACCTTGACACACTGCAGTGACGTCGCCGACTTTGCCCACCACCATCTGGCAGCAG 492

Qy   361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCTTC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCTTC 552

RESULT 12
HSGCSFR1
LOCUS   HSGCSFR1   1498 bp   mRNA   linear   PRI 21-MAR-1995
DEFINITION   Human mRNA for granulocyte colony-stimulating factor (G-CSF)
               (pBRV-2).
ACCESSION   X03655
VERSION   X03655.1 GI:31693
KEYWORDS   colony stimulating factor; signal peptide.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1498)
AUTHORS   Nagata,S., Tsuchiya,M., Asano,S., Yamamoto,O., Hirata,Y.,
               Kubota,N., Oheda,M., Nomura,H. and Yamazaki,T.
TITLE   The chromosomal gene structure and two mRNAs for human granulocyte
               colony-stimulating factor
JOURNAL   EMBO J. 5 (3), 575-581 (1986)
MEDLINE   86220137
PUBMED   2423327
COMMENT   Data kindly reviewed (19-JUN-1986) by S. Nagata.
FEATURES   Location/Qualifiers
     source   1..1498
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS   31..645
                     /codon_start=1
                     /product="G-CSF protein"
                     /protein_id="CAA27290.1"
                     /db_xref="GI:732764"
                     /db_xref="GOA:P09919"
                     /db_xref="SWISS-PROT:P09919"
                     /translation="MAGPATQSPMKLMALQLLLWHSALWTVQEATPLGPASSLPQSFL
                     LKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLA
                     GCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGMAPAL
                     QPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP"
     sig_peptide   31..120
                     /note="signal peptide (aa -30 to -1)"
     mat_peptide   121..642
                     /product="mature G-CSF protein (aa 1-174)"
     misc_feature   1481..1486
                     /note="polyadenylation signal"
     polyA_site   1498
                     /note="polyadenylation site"
ORIGIN

Query Match   100.0%; Score 523; DB 9; Length 1498;
Best Local Similarity 100.0%; Pred. No. 2.1e-86;
Matches   523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACCCCCCTGGGCCCTGCAGCTCCTGCCCAGAGCTCCTCTGCTCAAGTGCTTAGAGCAA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACCCCCCTGGGCCCTGCAGCTCCTGCCCAGAGCTCCTCTGCTCAAGTGCTTAGAGCAA 180

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGCAGCGGCTCCAGGAGAAGTCGTGTGCCACCTACAAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGAGGAAGATCCAGGGCGATGGCGCAGCGGCTCCAGGAGAAGTCGTGTGCCACCTACAAG 240

Qy   121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTTCTGGGCATCCCCTGGGCTCCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTTCTGGGCATCCCCTGGGCTCCC 300

Qy   421 GCCTCTGCTTCCAGGCGCCGGCAGGGGGTCCTAGTGCCTCCATCTGCAGAGCTTC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   553 GCCTCTGCTTCCAGGCGCCGGCAGGGGGTCCTAGTGCCTCCATCTGCAGAGCTTC 612

Qy   481 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 523
           |||||||||||||||||||||||||||||||||||||||||||
Db   613 CTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCT 655

Query Match   100.0%; Score 523; DB 6; Length 1521;
Best Local Similarity   100.0%; Pred. No. 2.1e-86;
Matches 523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ACCCCCTGGGCCCTGCCAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACCCCCTGGGCCCTGCCAGCTTCCTGCTCAAGTGCTTAGAGCAA 180

Qy     61 GTGAGGAAGATCAGGGCGATGGCGCAGCGTCCAGGAGAAGCTTGTGTGCCACCTACAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGAGGAAGATCAGGGCGATGGCGCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 240

Qy    121 CTGTGCCACCCCGAGAGCTGTGCTCGGACACTCTCTGGCATCCCTGGCTCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGTGCCACCCCGAGAGCTGGTGCTGTGGACACTCTCTGGGCATCCCTGGGCTCCC 300

Qy    181 CTGAGCAGCTGCCCCAGCCCCTGCTCTTGCGAGGTGCTGCCAGCCAACTCCATAGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGAGCAGCTGCCCCAGCCCCTGCTCTTGCGAGGTGCTGCCAGCCAACTCCATAGC 360

Qy    241 GGCCTTTTCCTCTACCAGGGGATCTCCCGAGTTGGGT 300
          |||||||||||||||||||||||||||||||||||||
Db    361 GGCCTTTTCCTCTACCAGGGGATCTCCCGAGTTGGGT 420

Qy    301 CCCACCTTGGACACACTGCAGTGGACGTCGCGTCGCGGACTTTGCCACCACCATCTGGCAGCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCCACCTTGGACACACTGCAGTGGACGTGCGCCCCAGCCCCACCAGGGGTGCCATGCCGGCTTC 480

Qy    361 ATGGAAGAACTGGGAATGGCCCCTGCCTAGTTGCCTCCATCTGCAGAGCTTC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGGAAGAACTGGGAATGGCCCCTGCCTAGTTGCCTCCATCTGCAGAGCTTC 540

Qy    421 GCCTCTGCTTTCCAGCCCGGGCCAGGAGGGGTCCTCCAGTTGCCCAGCCCT 523
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCCTCTGCTTTCCAGCCCGGGCCAGGAGGGGTCCTCCACGGTTGCCCAGCCCT 643
```

RESULT 13
E01219
LOCUS   E01219   1521 bp   RNA   linear   PAT 29-SEP-1997
DEFINITION   cDNA encoding human G-CSF.
ACCESSION   E01219
VERSION   E01219.1   GI:2169478
KEYWORDS   JP 1987129298-A/1.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1521)
AUTHORS   Yamazaki,T., Yamamoto,O., Hirata,Y., Sekimori,Y. and Osada,J.
TITLE   NOVEL POLYPEPTIDE
JOURNAL   Patent: JP 1987129298-A 1 11-JUN-1987;
             CHUGAI PHARMACEUT CO LTD
COMMENT   OS   Human
             PN   JP 1987129298-A/1
             PD   11-JUN-1987
             PF   02-DEC-1985 JP 1985269455
             PI   YAMAZAKI TATSUMI, YAMAMOTO OSAMI, HIRATA YUICHI, PI   SEKIMORI
             YASUO,
             PI   OSADA JUICHI
             PC   C07K13/00,C07H21/04,C12N15/00,C12P21/02;
             CC   strandedness: Double;
             CC   topology: Linear;
             CC   hypothetical: No;
             CC   anti-sense: No;
             CC   *source: cell_type=oral cavity tumor cells;
             CC   *source: cell_line=CHU-2;
             FH   Key   Location/Qualifiers
             FH
             FT   mRNA   1..1521
             FT   CDS   31..645
             FT   /product='human G-CSF'
             FT   sig_peptide   31..120
             FT   /note='human G-CSF prewsor'
             FT   mat_peptide   121..642
             FT   /note='human G-CSF'
             FT   5'UTR   1..30
             FT   3'UTR   645..1499
             FT   polyA_site   1500..1522
             FT   mRNA   1..1522.
ORIGIN

RESULT 14
I05333
LOCUS   I05333   1525 bp   DNA   linear   PAT 02-DEC-1994
DEFINITION   Sequence 11 from Patent EP 0256843.
ACCESSION   I05333
VERSION   I05333.1   GI:591091
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1525)
AUTHORS   Devlin,J.J., Devlin,P.E., Kawasaki,E.S. and Warren,M.K.
TITLE   Expression of g-csf and muteins thereof and their use
JOURNAL   Patent: EP 0256843-A1 11 24-FEB-1988;
FEATURES   Location/Qualifiers
     source   1..1525
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   100.0%; Score 523; DB 6; Length 1525;
Best Local Similarity   100.0%; Pred. No. 2.1e-86;
Matches 523; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ACCCCCTGGGCCCTGCCAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||
Db    133 ACCCCCTGGGCCCTGCCAGCTTCCTGCTCAAGTGCTTAGAGCAA 192

Qy     61 GTGAGGAAGATCAGGGCGATGGCGCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 GTGAGGAAGATCAGGGCGATGGCGCAGCGTCCAGGAGAAGCTGTGTGCCACCTACAAG 252
```

```
Qy  301 CCCACCTTGGACACACTGCAGCTGGACGTGCGCGACTTTGCCACCACCATCTGGCAGCAG 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  433 CCCACCTTGGACACACTGCAGCTGGACGTGCGCGACTTTGCCACCACCATCTGGCAGCAG 492

Qy  361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCTGCGAGCCCACCCAGGTGCCATGCCGGCCTTC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  493 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCTGCGAGCCCACCCAGGTGCCATGCCGGCCTTC 552

Qy  421 GCCTCTGCTTCCAGCGCCGCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  553 GCCTCTGCTTCCAGCGCCGCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 612

Qy  481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
        |||||||||||||||||||||||||||||||||||||||||||
Db  613 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 655


Search completed: October  9, 2004, 03:08:31
Job time : 2729.62 secs
```

```
Qy  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGACACTCTCTGGGCATCCCCTGGGCTCCCC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  253 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGACACTCTCTGGGCATCCCCTGGGCTCCCC 312

Qy  181 CTGAGCAGCTGCCCCAGCCAGCCCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  313 CTGAGCAGCTGCCCCAGCCAGCCCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 372

Qy  241 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  373 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 432

Qy  301 CCCACCTTGGACACACTGCAGCTGCTGCCGACGTGCGCCGACTTTGCCACCACCATCTGGCAGCAG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  433 CCCACCTTGGACACACTGCAGCTGCTGCCGACGTGCGCCGACTTTGCCACCACCATCTGGCAGCAG 492

Qy  361 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCTGCGAGCCCACCCAGGTGCCATGCCGGCCTTC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  493 ATGGAAGAACTGGGAATGGCCCCTGCCCTGCCCTGCGAGCCCACCCAGGTGCCATGCCGGCCTTC 552

Qy  421 GCCTCTGCTTCCAGCGCCGCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  553 GCCTCTGCTTCCAGCGCCGCGGCAGGAGGGGTCCTAGTTGCCTCCCATCTGCAGAGCTTC 612

Qy  481 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 523
        |||||||||||||||||||||||||||||||||||||||||||
Db  613 CTGGAGGTGTCGTACCGGCGTTCTACGCCACCTTGCCCAGCCCT 655
```

```
RESULT 15
LOCUS       I08669
I08669
LOCUS       I08669                  1525 bp    DNA        linear    PAT 02-DEC-1994
DEFINITION  Sequence 1 from Patent WO 8801297.
ACCESSION   I08669
VERSION     I08669.1  GI:58622
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1525)
  AUTHORS   Devlin,J.J., Kawasaki,E.S. and Warren,M.K.
  JOURNAL   Patent: WO 8801297-A 1 25-FEB-1988;
FEATURES             Location/Qualifiers
     source          1..1525
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match        100.0%;  Score 523;  DB 6;  Length 1525;
Best Local Similarity  100.0%;  Pred. No. 2.1e-86;
Matches  523;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCACCCGAGCTTCCTGCTCAAGTGCTTAGAGCAA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  133 ACCCCCTGGGCCCTGCCCAGCTCCCTGCCCCACCCGAGCTTCCTGCTCAAGTGCTTAGAGCAA 192

Qy   61 GTGAGGAAGATCCAGGGCGATGGCGAGGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  193 GTGAGGAAGATCCAGGGCGATGGCGAGGGCGCTCCAGGAGAAGCTGTGTGCCACCTACAAG 252

Qy  121 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGACACTCTCTGGGCATCCCCTGGGCTCCCC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  253 CTGTGCCACCCCGAGGAGCTGGTGCTGCTCGACACTCTCTGGGCATCCCCTGGGCTCCCC 312

Qy  181 CTGAGCAGCTGCCCCAGCCAGCCCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  313 CTGAGCAGCTGCCCCAGCCAGCCCCTGCAGTGGCAGGCTGCTTGAGCCAACTCCATAGC 372

Qy  241 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  373 GGCCTTTTCCTCTACCAGGGGCTCCTGCGAGGCCCTGGAAGGGATCTCCCCGAGTTGGGT 432
```

This Page Blank (uspto)

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2004, 00:53:40 ; Search time 1111.24 Seconds
(without alignments)
4837.123 Million cell updates/sec

Title: US-10-009-792C-26
Perfect score: 180
Sequence: 1 atgtttaagttaaaagaa.........tcctgctcaagtgcttagag 180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1493109276 residues

Total number of hits satisfying chosen parameters: 55025578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
    1: em_estba:*
    2: em_esthum:*
    3: em_estin:*
    4: em_estmu:*
    5: em_estov:*
    6: em_estpl:*
    7: em_estro:*
    8: em_htc:*
    9: gb_est1:*
   10: gb_est2:*
   11: gb_htc:*
   12: gb_est3:*
   13: gb_est4:*
   14: gb_est5:*
   15: em_estfun:*
   16: em_estom:*
   17: em_gss_hum:*
   18: em_gss_inv:*
   19: em_gss_pln:*
   20: em_gss_vrt:*
   21: em_gss_fun:*
   22: em_gss_mam:*
   23: em_gss_mus:*
   24: em_gss_pro:*
   25: em_gss_rod:*
   26: em_gss_phg:*
   27: em_gss_vrl:*
   28: gb_gss1:*
   29: gb_gss2:*

No.	Score	%			c	Accession	ID / Description
5	49	27.2	598	12		BG548320	BG548320 602575289
6	49	27.2	608	29		AY414901	AY414901 Pan trogl
7	49	27.2	622	14		CD013926	CD013926 9013855
8	49	27.2	624	29		AY414900	AY414900 Homo sapi
9	49	27.2	789	14		CD013925	CD013925 9013939
10	49	27.2	802	14		CA489740	CA489740 AGENCOURT
11	49	27.2	817	12		BM009358	BM009358 603629823
12	49	27.2	897	14		CD013924	CD013924 9013889
13	49	27.2	906	14		CA487627	CA487627 AGENCOURT
14	49	27.2	948	12		BI822673	BI822673 603036006
15	49	27.2	983	12		BM557421	BM557421 AGENCOURT
16	49	27.2	986	14		CA487516	CA487516 AGENCOURT
17	49	27.2	990	12		BM423896	BM423896 AGENCOURT
18	49	27.2	993	12		BM915836	BM915836 AGENCOURT
19	49	27.2	1012	12		BM556479	BM556479 AGENCOURT
20	49	27.2	1080	12		BM553432	BM553432 AGENCOURT
21	49	27.2	1096	12		BM923410	BM923410 AGENCOURT
22	49	27.2	1123	12		BM554035	BM554035 AGENCOURT
23	49	27.2	1141	12		BM906188	BM906188 AGENCOURT
24	49	27.2	1445	12		BM906367	BM906367 AGENCOURT
25	47.4	26.3	368	12		BI961002	BI961002 MONO1_3_H
26	47.4	26.3	548	12		BI960812	BI960812 MONO1_1_B
27	47.4	26.3	570	12		BI961242	BI961242 MONO1_7_G
28	44.2	24.6	509	12		BM256261	BM256261 518592_MA
29	44.2	24.6	513	10		BE480590	BE480590 165716_BA
30	44.2	24.6	1375	12		BM554037	BM554037 AGENCOURT
31	36.4	20.2	777	12		BG571954	BG571954 602592408
32	34.6	19.2	888	13	c	BQ716545	BQ716545 AGENCOURT
33	34.2	19.0	887	29	c	CNS02GVH	AL196838 Tetraodon
34	33.8	18.8	754	29		CNS004TE	AL055711 Drosophil
35	33.4	18.6	829	12		BM009247	BM009247 603629663
36	33.2	18.4	675	28	c	CC100458	CC100458 CSU-K34.1
37	33	18.3	538	14		CA621905	CA621905 wlln.pk00
38	33	18.3	1005	29		CG998971	CG998971 ZMMBBb052
39	33	18.3	1100	12	c	BI079593	BI079593 602876413
40	32.8	18.2	854	29	c	CG674466	CG674466 trs2198 t
41	32.6	18.1	489	29	c	AQ721853	AQ721853 HS_560_B
42	32.6	18.1	924	29	c	CNS047GV	CNS047GV Tetraodon
43	32.4	18.0	1216	10		BF137044	BF137044 601783256
44	32.2	17.9	454	13	c	BX565066	BX565066 BX565066
45	32.2	17.9	473	13	c	BX557098	BX557098 BX557098

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	52	28.9	346	14	CB705562	CB705562 AMGNNUC:T
2	49	27.2	314	14	CB136760	CB136760 K-EST0189
3	49	27.2	554	14	CB126975	CB126975 K-EST0176
4	49	27.2	579	14	CB127794	CB127794 K-EST0177

ALIGNMENTS

RESULT 1
CB705562
LOCUS   CB705562   346 bp   mRNA   linear   EST 10-APR-2003
DEFINITION AMGNNUC:TRCX1-00003-C7-A trcx1 (10261) Rattus norvegicus cDNA clone
            trcx1-00003-c7 5', mRNA sequence.
ACCESSION   CB705562
VERSION   CB705562.1 GI:29762710
KEYWORDS   EST.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 346)
AUTHORS   Amgen EST Program.
TITLE   Amgen Rat EST Program
JOURNAL   Unpublished (2003)
COMMENT   Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00003 row: c column: 7.
FEATURES   Location/Qualifiers
     source   1..346
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"

```
                /clone="trcx1-00003-c7"
                /tissue_type="choroid plexus brain"
                /clone_lib="trcx1 (10261)"
                /note="Vector; pcAv; Site_1: Sall; Site_2: NotI; choroid
                plexus brain region"

ORIGIN

        Query Match     28.9%;   Score 52;   DB 14;   Length 346;
        Best Local Similarity   72.8%;   Pred. No. 3.4e-05;
        Matches   67;   Conservative   0;   Mismatches   25;   Indels   0;   Gaps   0;

Qy   89  GCCGCCACCATCACCATACCATATCGAGGAAGGACTCCGTTAGGTCCAGCCAGCTCCC 148
            |||| |||||  ||| |||||||||||| |||||||||||| |||||||
Db   28  GCCCCGGACTCCAAGCTCCATGACCAGCCAGTCTAGAACCCCCCGGGCCCTGCCAGCTCCC 87

Qy   149 TGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
            ||||||||||||||||||||||||||||||
Db   88  TGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 119

RESULT 2
CB136760
LOCUS       CB136760         314 bp    mRNA        linear    EST 29-JAN-2003
DEFINITION  K-EST0189338 L5HLK1 Homo sapiens cDNA clone L5HLK1-42-H10 5', mRNA
            sequence.
ACCESSION   CB136760
VERSION     CB136760.1   GI:28104242
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 314)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: Yongsung@mail.kribb.re.kr
            Plate: 42  row: H  column: 10
            High quality sequence stop: 314.
FEATURES             Location/Qualifiers
     source          1..314
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L5HLK1-42-H10"
                     /sex="M"
                     /cell_line="HLK-1"
                     /lab_host="Top10F'"
                     /clone_lib="L5HLK1"
                     /note="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tobacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."


ORIGIN

        Query Match     27.2%;   Score 49;   DB 14;   Length 314;
        Best Local Similarity   84.6%;   Pred. No. 0.00028;
        Matches   55;   Conservative   0;   Mismatches   10;   Indels   0;   Gaps   0;

Qy   116 AGGGAAGGACTCCGTTAGGTCCAGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
            |||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||
Db   118 AGGAAGCCACCCCCTGGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 177

Qy   176 TAGAG 180
            |||||
Db   178 TAGAG 182

RESULT 3
CB126975
LOCUS       CB126975         554 bp    mRNA        linear    EST 29-JAN-2003
DEFINITION  K-EST0176151 C1SNU17 Homo sapiens cDNA clone C1SNU17-14-A07 5',
            mRNA sequence.
ACCESSION   CB126975
VERSION     CB126975.1   GI:28089170
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 14  row: A  column: 07
            High quality sequence stop: 554.
FEATURES             Location/Qualifiers
     source          1..554
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="C1SNU17-14-A07"
                     /sex="F"
                     /tissue_type="Uterine"
                     /cell_type="Epithelial"
                     /cell_line="SNU-17"
                     /lab_host="Top10F'"
                     /clone_lib="C1SNU17"
                     /note="Organ: Cervix; Vector: pCNS-D2; Site 1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tobacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."

ORIGIN

        Query Match     27.2%;   Score 49;   DB 14;   Length 554;
```

Best Local Similarity   84.6%;   Pred. No. 0.00034;
Matches 55;   Conservative   0;   Mismatches 10;   Indels   0;   Gaps   0;

```
Qy   116 AGGGAAGGAGGACTCCGTTAGGTCCAGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
         |||| ||  ||  |||||||||| |||||| |||||||||||||||||||||||||||||
Db   121 AGGAAGCCACCCCCCTGGGCCTGCCAGCTCCTGCCAGAGCTTCCTGCTCAAGTGCT 180

Qy   176 TAGAG 180
         |||||
Db   181 TAGAG 185
```

RESULT 4
CB127794
LOCUS   CB127794   579 bp   mRNA   linear   EST 29-JAN-2003
DEFINITION K-EST0177131 C1SNU17 Homo sapiens cDNA clone C1SNU17-10-C09 5',
            mRNA sequence.
ACCESSION   CB127794
VERSION   CB127794.1 GI:28090473
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 579)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE   21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 10 row: C column: 09
            High quality sequence stop: 579.
FEATURES   Location/Qualifiers
     source   1..579
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="C1SNU17-10-C09"
                     /sex="F"
                     /tissue_type="Uterine"
                     /cell_type="Epithelial"
                     /cell_line="SNU-17"
                     /clone_lib="C1SNU17"
                     /note="Organ: Cervix; Vector: pCNS-D2; Site 1: EcoRI;
                     Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN

Query Match   27.2%;   Score 49;   DB 14;   Length 579;
Best Local Similarity   84.6%;   Pred. No. 0.00035;
Matches 55;   Conservative   0;   Mismatches 10;   Indels   0;   Gaps   0;

```
Qy   116 AGGGAAGGAGGACTCCGTTAGGTCCAGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
         |||| ||  ||  |||||||||| |||||| |||||||||||||||||||||||||||||
Db   197 AGGAAGCCACCCCCCTGGGCCTGCCAGCTCCTGCCAGAGCTTCCTGCTCAAGTGCT 256

Qy   176 TAGAG 180
         |||||
Db   257 TAGAG 261
```

RESULT 5
BG548320
LOCUS   BG548320   598 bp   mRNA   linear   EST 04-APR-2001
DEFINITION 602575289F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703159 5',
            mRNA sequence.
ACCESSION   BG548320
VERSION   BG548320.1 GI:13546985
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 598)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1541 row: b column: 24
            High quality sequence stop: 597.
FEATURES   Location/Qualifiers
     source   1..598
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4703159"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
                     SfiI (ggccgcctcggcc); Site 2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Query Match   27.2%;   Score 49;   DB 12;   Length 598;
Best Local Similarity   84.6%;   Pred. No. 0.00035;
Matches 55;   Conservative   0;   Mismatches 10;   Indels   0;   Gaps   0;

```
Qy   116 AGGGAAGGAGGACTCCGTTAGGTCCAGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
         |||| ||  ||  |||||||||| |||||| |||||||||||||||||||||||||||||
Db    80 AGGAAGCCACCCCCCTGGGCCTGCCAGCTCCTGCCAGAGCTTCCTGCTCAAGTGCT 139

Qy   176 TAGAG 180
         |||||
Db   140 TAGAG 144
```

RESULT 6
AY414901
LOCUS   AY414901   608 bp   DNA   linear   GSS 17-DEC-2003
DEFINITION Pan troglodytes CSF3 gene, VIRTUAL TRANSCRIPT, partial sequence,

```
FEATURES             Location/Qualifiers
     source          1..622
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Single gene library"
                     /note="Vector: pDrive Cloning Vector; RT-PCR was performed
                     using gene-specific primers flanking the open-reading
                     frame. PCR products were subcloned into pDrive Cloning
                     Vector and sequenced completely using M13 forward and
                     reverse primers. Sequencing gaps were closed by
                     re-sequencig using primers flanking the gapped areas."
ORIGIN

Query Match        27.2%;  Score 49;  DB 14;  Length 622;
Best Local Similarity  84.6%;  Pred. No. 0.00036;
Matches  55;  Conservative  0;  Mismatches  10;  Indels  0;  Gaps  0;

Qy  116 AGGGAAGGACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
        ||| |||| ||||| || |||||||||||||||||||||||||||||||||||||||||
Db   91 AGGAAGCCACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 150

Qy  176 TAGAG 180
        |||||
Db  151 TAGAG 155
```

RESULT 8
AY414900

```
LOCUS       AY414900                624 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Homo sapiens CSF3 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY414900
VERSION     AY414900.1  GI:39770859
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 624)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
   TITLE     Direct Submission
   JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..624
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
      gene           <1..>624
                     /gene="CSF3"
                     /locus_tag="HCM5378"
ORIGIN

Query Match        27.2%;  Score 49;  DB 29;  Length 624;
Best Local Similarity  84.6%;  Pred. No. 0.00036;
Matches  55;  Conservative  0;  Mismatches  10;  Indels  0;  Gaps  0;

Qy  116 AGGGAAGGACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
```

```
ACCESSION   AY414901
VERSION     AY414901.1  GI:39770860
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 608)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 608)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
   TITLE     Direct Submission
   JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..608
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
      gene           <1..>608
                     /gene="CSF3"
                     /locus_tag="HCM5378"
ORIGIN

Query Match        27.2%;  Score 49;  DB 29;  Length 608;
Best Local Similarity  84.6%;  Pred. No. 0.00035;
Matches  55;  Conservative  0;  Mismatches  10;  Indels  0;  Gaps  0;

Qy  116 AGGGAAGGACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
        ||| |||| ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   67 AGGAAGCCACCCCCCTGGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 126

Qy  176 TAGAG 180
        |||||
Db  127 TAGAG 131
```

RESULT 7
CD013926

```
LOCUS       CD013926                622 bp    mRNA    linear   EST 21-OCT-2003
DEFINITION  90138855 Single gene library Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD013926
VERSION     CD013926.1  GI:37777456
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 622)
  AUTHORS   Jin,P., Fu,G.K. Wilson,A.D., Yang,J., Chien,D., Hawkins,P.R.,
            Au-Young,J. and Stuve,L.L.
  TITLE     PCR isolation and cloning of novel splice variant mRNAs from known
            drug target genes
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Jin, P.
            Incyte Corporation
            3160 Porter Drive, Palo Alto, CA 94304, USA
            Tel: 650 621 8639
            Fax: 650 621 8965
            Email: pjin@incyte.com.
```

```
Db      83 AGGAAGCCACCCCCCTGGCCCTGCCAGCTCCCTGCCCAGAGCTTCCTGCTCAAGTGCT 142
           ||||  |||  || |  ||  | ||||||||||||||||||||||||||||
Qy     176 TAGAG 180
           |||||
Db     143 TAGAG 147

RESULT 9
CD013925
LOCUS       CD013925
DEFINITION  9013939 Single gene library Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD013925
VERSION     CD013925.1  GI:37777455
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 789)
AUTHORS     Jin,P., Fu,G.K., Wilson,A.D., Yang,J., Chien,D., Hawkins,P.R.,
            Au-Young,J. and Stuve,L.L.
TITLE       PCR isolation and cloning of novel splice variant mRNAs from known
            drug target genes
JOURNAL     Unpublished (2003)
COMMENT     Contact: Jin,P.
            Incyte Corporation
            3160 Porter Drive, Palo Alto, CA 94304, USA
            Tel: 650 621 8639
            Fax: 650 621 8965
            Email: pjin@incyte.com.
FEATURES             Location/Qualifiers
     source          1..789
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Single gene library"
                     /note="Vector: pDrive Cloning Vector; RT-PCR was performed
                     using gene-specific primers flanking the open-reading
                     frame. PCR products were subcloned into pDrive Cloning
                     Vector and sequenced completely using M13 forward and
                     reverse primers. Sequencing gaps were closed by
                     re-sequencig using primers flanking the gapped areas."
ORIGIN

Query Match             27.2%;  Score 49;  DB 14;  Length 789;
Best Local Similarity   84.6%;  Pred. No. 0.00039;
Matches    55;  Conservative    0;  Mismatches  10;  Indels    0;  Gaps     0;

Qy     116 AGGAAGGAGACTCCGTTAGGTCCAGCCAGCGACTCCTGCCCCCAGAGCTTCCTGCTCAAGTGCT 175
           ||  |  |  ||  || ||  ||||||||||||||||||||||||||||||||||||||||
Db     267 AGGAAGCCACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCCAGAGCTTCCTGCTCAAGTGCT 326
Qy     176 TAGAG 180
           |||||
Db     327 TAGAG 331

RESULT 10
CA489740
LOCUS       CA489740              802 bp    mRNA    linear   EST 14-NOV-2002
DEFINITION  AGENCOURT_10810586 MAPcL Homo sapiens cDNA clone IMAGE:6722223 5',
            mRNA sequence.
ACCESSION   CA489740
VERSION     CA489740.1  GI:24952531
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 802)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Kristi A. Egland, Ira Pastan
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14284  row: j  column: 15
            High quality sequence stop: 249.
FEATURES             Location/Qualifiers
     source          1..802
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6722223"
                     /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
                     hTERT-HME1, LNCaP"
                     /lab_host="EMDH10B"
                     /clone_lib="MAPcL"
                     /note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
                     Subtracted with brain, liver, lung, kidney and muscle.
                     Directionally cloned. Priming method: oligo-dT. Average
                     insert size: 1800 bp. Library amplification: 26,000 fold.
                     Kristi A. Egland, James J. Vincent, Robert Strausberg,
                     Bungkook Lee & Ira Pastan: Discovery of new breast
                     cancer genes encoding membrane and secreted proteins.
                     Manuscript submitted."
ORIGIN

Query Match             27.2%;  Score 49;  DB 14;  Length 802;
Best Local Similarity   84.6%;  Pred. No. 0.00039;
Matches    55;  Conservative    0;  Mismatches  10;  Indels    0;  Gaps     0;

Qy     116 AGGAAGGAGACTCCGTTAGGTCCAGCCAGCGACTCCTGCCCCCAGAGCTTCCTGCTCAAGTGCT 175
           ||  |  |  ||  || ||  ||||||||||||||||||||||||||||||||||||||||
Db     107 AGGAAGCCACCCCCCTGGCCCTGCCAGCTCCCTGCCCCCAGAGCTTCCTGCTCAAGTGCT 166
Qy     176 TAGAG 180
           |||||
Db     167 TAGAG 171

RESULT 11
BM009358
LOCUS       BM009358              817 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  6036298223F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5443757 5',
            mRNA sequence.
ACCESSION   BM009358
VERSION     BM009358.1  GI:16523712
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 817)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1921  row: m  column: 06
            High quality sequence stop: 745.
FEATURES             Location/Qualifiers
```

```
     source          1..817
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5443757"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"
                     /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."

ORIGIN

Query Match        27.2%;    Score 49;   DB 12;   Length 817;
Best Local Similarity  84.6%;   Pred. No. 0.00039;
Matches   55;    Conservative   0;    Mismatches 10;    Indels   0;    Gaps    0;

Qy  116 AGGGAAGGACTCCGTTAGTCCAGCCAGTTCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
        ||| ||||||| ||||     ||||||||||||||||||||||||||||||||||||||
Db  270 AGGAAGCCACCCCCTGGGCCCTGCCAGTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 329

Qy  176 TAGAG 180
        |||||
Db  330 TAGAG 334


RESULT 12
CD013924
LOCUS       CD013924         897 bp    mRNA    linear   EST 21-OCT-2003
DEFINITION  90138839 Single gene library Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD013924
VERSION     CD013924.1  GI:37777454
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 897)
AUTHORS     Jin,P., Fu,G.K., Wilson,A.D., Yang,J., Chien,D., Hawkins,P.R.,
            Au-Young,J. and Stuve,L.L.
TITLE       PCR isolation and cloning of novel splice variant mRNAs from known
            drug target genes
JOURNAL     Unpublished (2003)
COMMENT     Contact: Jin, P.
            Incyte Corporation
            3160 Porter Drive, Palo Alto, CA 94304, USA
            Tel: 650 621 8639
            Fax: 650 621 8965
            Email: pjin@incyte.com.
FEATURES             Location/Qualifiers
     source          1..897
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Single gene library"
                     /note="Vector: pDrive Cloning Vector; RT-PCR was performed
                     using gene-specific primers flanking the open-reading
                     frame. PCR products were subcloned into pDrive Cloning
                     Vector and sequenced completely using M13 forward and
                     reverse primers. Sequencing gaps were closed by
                     re-sequencig using primers flanking the gapped areas."

ORIGIN

Query Match        27.2%;    Score 49;   DB 14;   Length 897;
Best Local Similarity  84.6%;   Pred. No. 0.00041;
Matches   55;    Conservative   0;    Mismatches 10;    Indels   0;    Gaps    0;

Qy  116 AGGGAAGGACTCCGTTAGTCCAGCCAGTTCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
        ||| ||||| |||||     ||||||||||||||||||||||||||||||||||||||
Db  267 AGGAAGCCACCCCCTGGGCCCTGCCAGTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 326

Qy  176 TAGAG 180
        |||||
Db  327 TAGAG 331


RESULT 13
CA487627
LOCUS       CA487627         906 bp    mRNA    linear   EST 14-NOV-2002
DEFINITION  AGENCOURT_10808790 MAPcL Homo sapiens cDNA clone IMAGE:6719127 5',
            mRNA sequence.
ACCESSION   CA487627
VERSION     CA487627.1  GI:24947489
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 906)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Kristi A. Egland, Ira Pastan
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14276 row: i  column: 15
            High quality sequence stop: 712.
FEATURES             Location/Qualifiers
     source          1..906
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6719127"
                     /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
                     hTERT-HME1, LNCaP"
                     /lab_host="EMDH10B"
                     /clone_lib="MAPcL"
                     /note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
                     Subtracted with brain, liver, lung, kidney and muscle.
                     Directionally cloned. Priming method: oligo-dT. Average
                     insert size: 1800 bp. Library amplification: 26,000 fold.
                     Kristi A. Egland, James J. Vincent, Robert Strausberg,
                     Bungkook Lee & Ira Pastan: Discovery of new breast
                     cancer genes encoding membrane and secreted proteins.
                     Manuscript submitted."

ORIGIN

Query Match        27.2%;    Score 49;   DB 14;   Length 906;
Best Local Similarity  84.6%;   Pred. No. 0.00041;
Matches   55;    Conservative   0;    Mismatches 10;    Indels   0;    Gaps    0;

Qy  116 AGGGAAGGACTCCGTTAGTCCAGCCAGTTCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
        ||| ||||| |||||     ||||||||||||||||||||||||||||||||||||||
Db   72 AGGAAGCCACCCCCTGGGCCCTGCCAGTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 131

Qy  176 TAGAG 180
        |||||
Db  132 TAGAG 136


RESULT 14
BI822673
LOCUS       BI822673         948 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603036006F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5177209 5',
            mRNA sequence.
```

```
ACCESSION   BI822673
VERSION     BI822673.1  GI:15934223
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1441 row: k  column: 02
            High quality sequence stop: 845.
FEATURES    Location/Qualifiers
     source  1..948
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5177209"
             /lab_host="DH10B"
             /clone_lib="NIH_MGC 115"
             /note="Organ: pooled brain, lung, testis; Vector:
             pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
             source anonymous pool of 6 male_brains, age range 23-27; 1
             male lung, age 27, and 1 male testis, age 69. Library is
             oligo-dT primed and directionally cloned (EcoRV site is
             destroyed upon cloning). Average insert size 1.8 kb,
             insert size range 1-3 kb. Library is normalized and
             enriched for full-length clones and was constructed by C.
             Gruber (Invitrogen). Research Genetics tracking code
             021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match            27.2%;   Score 49;   DB 12;   Length 948;
Best Local Similarity  84.6%;   Pred. No. 0.00042;
Matches  55;   Conservative  0;   Mismatches  10;   Indels  0;   Gaps  0;

Qy  116 AGGAAGGACTCCGTTAGGTCCAGCCAGTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
        ||||| |||| || ||||||||||||||||||||||||||||||||||||||||||||
Db  287 AGGAAGCCACCCCCCTGGGCCCTGCAGCTTCCTGCAGCTTCCTGCTCAAGTGCT 346

Qy  176 TAGAG 180
        |||||
Db  347 TAGAG 351


RESULT 15
BM557421
LOCUS       BM557421     983 bp    mRNA      linear  EST 20-FEB-2002
DEFINITION  AGENCOURT_6578800 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466747
            5', mRNA sequence.
ACCESSION   BM557421
VERSION     BM557421.1  GI:18799394
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 983)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1967 row: k  column: 04
            High quality sequence stop: 593.
FEATURES    Location/Qualifiers
     source  1..983
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5466747"
             /tissue_type="amelanotic melanoma, cell line"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC 41"
             /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
             EcoRI; cDNA made by oligo-dT priming. Directionally cloned
             into EcoRI/XhoI sites using the following 5' adaptor:
             GGCACGAG(G). Library constructed by Ling Hong in the
             laboratory of Gerald M. Rubin (University of California,
             Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
             Superscript II RT (Life Technologies). Note: this is a
             NIH_MGC Library."

ORIGIN

Query Match            27.2%;   Score 49;   DB 12;   Length 983;
Best Local Similarity  84.6%;   Pred. No. 0.00042;
Matches  55;   Conservative  0;   Mismatches  10;   Indels  0;   Gaps  0;

Qy  116 AGGAAGGACTCCGTTAGGTCCAGCCAGTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCT 175
        ||||| |||| || |||||||||||||||||||||||||||||||||||||||||||
Db  137 AGGAAGCCACCCCCCTGGGCCCTGCAGCTTCCTGCAGCTTCCTGCTCAAGTGCT 196

Qy  176 TAGAG 180
        |||||
Db  197 TAGAG 201


Search completed: October  9, 2004, 04:22:29
Job time : 1114.24 secs
```

This Page Blank (uspto)

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    October 9, 2004, 03:08:44 ; Search time 154.651 Seconds
                                   (without alignments)
                                   5900.506 Million cell updates/sec

Title:       US-10-009-792C-26
Perfect score: 180
Sequence:    1 atgtttaagttaaaagaa........tcctgctcaagtgcttagag 180

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched:    3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters:    6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

Database :   Published Applications NA:*
          1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
          2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
          3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
          4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
          5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
          6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
          7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
          8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
          9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
         10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
         11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
         12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
         13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
         14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
         15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
         16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
         17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
         18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
         19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	180	100.0	180	15	US-10-009-792A-26	Sequence 26, Appl
2	86	47.8	135	15	US-10-009-792A-24	Sequence 24, Appl
3	79	43.9	645	15	US-10-237-386-10	Sequence 10, Appl
4	72.8	40.4	657	15	US-10-237-386-11	Sequence 11, Appl
5	50.6	28.1	546	15	US-10-083-446-177	Sequence 177, Appl
6	50.6	28.1	546	15	US-10-083-446-178	Sequence 178, Appl
7	50.6	28.1	906	15	US-10-083-446-53	Sequence 53, Appl
8	50.6	28.1	921	15	US-10-083-446-72	Sequence 72, Appl
9	50.6	28.1	921	15	US-10-083-446-75	Sequence 75, Appl
10	50.6	28.1	921	15	US-10-083-446-78	Sequence 78, Appl
11	50.6	28.1	921	15	US-10-083-446-84	Sequence 84, Appl
12	50.6	28.1	966	15	US-10-083-446-68	Sequence 68, Appl
13	50.6	28.1	966	15	US-10-083-446-70	Sequence 70, Appl
14	50.6	28.1	966	15	US-10-083-446-71	Sequence 71, Appl
15	50.6	28.1	966	15	US-10-083-446-73	Sequence 73, Appl
16	50.6	28.1	966	15	US-10-083-446-77	Sequence 77, Appl
17	50.6	28.1	966	15	US-10-083-446-79	Sequence 79, Appl
18	50.6	28.1	1017	15	US-10-083-446-67	Sequence 67, Appl
19	50.6	28.1	1047	15	US-10-083-446-74	Sequence 74, Appl
20	50.6	28.1	1047	15	US-10-083-446-76	Sequence 76, Appl
21	49	27.2	507	15	US-10-009-792A-17	Sequence 17, Appl
22	49	27.2	531	17	US-10-695-584A-289	Sequence 289, App
23	49	27.2	531	17	US-10-695-584A-290	Sequence 290, App
24	49	27.2	531	17	US-10-695-584A-292	Sequence 292, App
25	49	27.2	531	17	US-10-695-584A-293	Sequence 293, App
26	49	27.2	531	17	US-10-695-584A-295	Sequence 295, App
27	49	27.2	531	17	US-10-695-584A-296	Sequence 296, App
28	49	27.2	531	17	US-10-695-584A-297	Sequence 297, App
29	49	27.2	531	17	US-10-695-584A-298	Sequence 298, App
30	49	27.2	531	17	US-10-695-584A-299	Sequence 299, App
31	49	27.2	615	15	US-10-695-584A-313	Sequence 313, App
32	49	27.2	531	17	US-10-009-792A-18	Sequence 18, Appl
33	49	27.2	630	13	US-10-609-346-17	Sequence 17, Appl
34	49	27.2	975	17	US-10-695-584A-263	Sequence 263, App
35	49	27.2	975	17	US-10-695-584A-264	Sequence 264, App
36	49	27.2	975	17	US-10-695-584A-265	Sequence 265, App
37	49	27.2	975	17	US-10-695-584A-266	Sequence 266, App
38	49	27.2	975	17	US-10-695-584A-267	Sequence 267, App
39	49	27.2	975	17	US-10-695-584A-268	Sequence 268, App
40	49	27.2	975	17	US-10-695-584A-269	Sequence 269, App
41	49	27.2	975	17	US-10-695-584A-270	Sequence 270, App
42	49	27.2	1365	10	US-09-968-362-21	Sequence 21, Appl
43	49	27.2	1368	10	US-09-968-362-17	Sequence 17, Appl
44	49	27.2	1371	10	US-09-968-362-19	Sequence 19, Appl
45	49	27.2	1498	12	US-10-447-315-18	Sequence 18, Appl

ALIGNMENTS

```
RESULT 1
US-10-009-792A-26
; Sequence 26, Application US/10009792A
; Publication No. US2003015304901
; GENERAL INFORMATION:
; APPLICANT: LEE, Sang-Yup
; APPLICANT: JEONG, Ki-Jun
; TITLE OF INVENTION: ESCHERICHIA COLI STRAIN SECRETING HUMAN
; TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR (G-CSF)
; FILE REFERENCE: HYLEE60.001APC
; CURRENT APPLICATION NUMBER: US/10/009,792A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: PCT/KR01/00549
; PRIOR APPLICATION NUMBER: 2001-03-31
; PRIOR APPLICATION NUMBER: KR 10-2000-0017052
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 26
;   LENGTH: 180
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-009-792A-26

  Query Match         100.0%;  Score 180;  DB 15;  Length 180;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-53;
  Matches  180;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGTTTAAGTTAAAAGAAATTCTTAGTGGGATTAACGGAGCAGCTTTCATGAGTATCAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGTTTAAGTTAAAAGAAATTCTTAGTGGGATTAACGGAGCAGCTTTCATGAGTATCAGC 60

Qy     61 ATGTTTCTGCAACCGGCCTCTGCAGCTGCCGCCGCACCATCACCATATCGAGGGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGTTTCTGCAACCGGCCTCTGCAGCTGCCGCCGCACCATCACCATATCGAGGGA 120

Qy    121 AGGACTCCGTTAGTCCAGCCAGCTCCCTGCCCCAGAGCTTCGCTCAAGTGCTTAGAG 180
```

Matches   94;   Conservative   0;   Mismatches   25;   Indels   0;   Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTTAATGAGTATTAGC 63

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCGCACCATCACCATATCGAGGG 119
            ||||||||||||||||||| ||||||| | |||| || ||| ||||
Db   64 TTGTTTTCGGCAACCGCCTCTGCGCTACTGGCAACTTGGCAAATTGGACTGATGG 122

RESULT 4
US-10-237-386-11
; Sequence 11, Application US/10237386
; Publication No. US2003018095A1
; GENERAL INFORMATION:
;   APPLICANT: Danisco A/S
;   APPLICANT: Sibbesen, Ole
;   APPLICANT: Sorensen, Jens
;   TITLE OF INVENTION: Xylanse Variants Having Altered Sensitivity to Xylanase Inhibitor
;   FILE REFERENCE: 674509-2046
;   CURRENT APPLICATION NUMBER: US/10/237,386
;   CURRENT FILING DATE: 2002-12-06
;   PRIOR APPLICATION NUMBER: PCT/IB01/00426
;   PRIOR FILING DATE: 2001-03-08
;   PRIOR APPLICATION NUMBER: GB 0005585.5
;   PRIOR FILING DATE: 2000-03-08
;   PRIOR APPLICATION NUMBER: GB 0015751.1
;   PRIOR FILING DATE: 2000-06-27
;   NUMBER OF SEQ ID NOS: 66
;   SOFTWARE: PatentIn version 3.0
SEQ ID NO 11
;   LENGTH: 657
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: B. subtilis xylanase sequence with added restriction site
US-10-237-386-11

Query Match   40.4%;   Score 72.8;   DB 15;   Length 657;
Best Local Similarity   91.7%;   Pred. No. 3.2e-15;
Matches   77;   Conservative   0;   Mismatches   7;   Indels   0;   Gaps   0;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
            ||||||||||||||||||||||||||||||| || ||||| ||||||| || ||||| |||
Db   4 ATGTTTAAGTTTAAAAAGAAATTCTTAGTTGGATTATCGGCAGCTTTAATGAGTATTAGC 63

Qy   61 ATGTTTCTGCAACCGCCTCTGCA 84
            ||||| ||||||||||||||||||
Db   64 TTGTTTTCGGCAACCGCCTCTGCA 87

RESULT 5
US-10-083-446-177
; Sequence 177, Application US/10083446
; Publication No. US2003018579 0A1
; GENERAL INFORMATION:
;   APPLICANT: Abrams, Mark A.
;   Bauer, S. C.
;   Braford-Goldberg, Sarah R.
;   Caparon, Maire H.
;   Easton, Alan M.
;   Klein, Barbara K.
;   McKearn, John P.
;   Olins, Peter O.
;   Paik, Kunnan
;   Thomas, John W.
;   TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
;   Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
;   NUMBER OF SEQUENCES: 197
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
;   Corporate Patent Dept., Mail Zone O4E

Db   121 AGGACTCCGTTAGTCCAGCCAGCGCTCCCTGCCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2
US-10-009-792A-24
; Sequence 24, Application US/10009792A
; Publication No. US2003015304 9A1
; GENERAL INFORMATION:
;   APPLICANT: LEE, Sang-Yup
;   APPLICANT: JEONG, Ki-Jun
;   TITLE OF INVENTION: ESCHERICHIA COLI STRAIN SECRETING HUMAN
;   TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR (G-CSF)
;   FILE REFERENCE: HYLEE60.001APC
;   CURRENT APPLICATION NUMBER: US/10/009,792A
;   CURRENT FILING DATE: 2002-10-29
;   PRIOR APPLICATION NUMBER: PCT/KR01/00549
;   PRIOR FILING DATE: 2001-03-31
;   PRIOR APPLICATION NUMBER: KR 10-2000-0017052
;   PRIOR FILING DATE: 2000-03-31
;   NUMBER OF SEQ ID NOS: 27
;   SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 24
;   LENGTH: 135
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-009-792A-24

Query Match   47.8%;   Score 86;   DB 15;   Length 135;
Best Local Similarity   77.6%;   Pred. No. 3.1e-20;
Matches   135;   Conservative   0;   Mismatches   0;   Indels   39;   Gaps   1;

Qy   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTTTAAGTTTAAAAAGAAATTCTTAGTGGGATTAACGGCAGCTTTCATGAGTATCAGC 60

Qy   61 ATGTTTCTGCAACCGCCTCTGCAGCTGGCCGCACCATCACCATATCGAGGGA 120
            |||||||||||||||||||||||||
Db   61 ATGTTTCTGCAACCGCCTCTGCA-------------------------------------- 84

Qy   121 AGGACTCCGTTAGTCCAGCCAGCGCTCCCTGCCCCAGCTTCCTGCTCCAAGTGC 174
                         ||||||||||||||||||||||||||||||||||||||||||||
Db   85 ---ACTCCGTTAGTCCAGCCAGCGCTCCCTGCCCCAGCTTCCTGCTCCAAGTGC 135

RESULT 3
US-10-237-386-10
; Sequence 10, Application US/10237386
; Publication No. US2003018095A1
; GENERAL INFORMATION:
;   APPLICANT: Danisco A/S
;   APPLICANT: Sibbesen, Ole
;   APPLICANT: Sorensen, Jens
;   TITLE OF INVENTION: Xylanse Variants Having Altered Sensitivity to Xylanase Inhibitor
;   FILE REFERENCE: 674509-2046
;   CURRENT APPLICATION NUMBER: US/10/237,386
;   CURRENT FILING DATE: 2002-12-06
;   PRIOR APPLICATION NUMBER: PCT/IB01/00426
;   PRIOR FILING DATE: 2001-03-08
;   PRIOR APPLICATION NUMBER: GB 0005585.5
;   PRIOR FILING DATE: 2000-03-08
;   PRIOR APPLICATION NUMBER: GB 0015751.1
;   PRIOR FILING DATE: 2000-06-27
;   NUMBER OF SEQ ID NOS: 66
;   SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
;   LENGTH: 645
;   TYPE: DNA
;   ORGANISM: Bacillus subtilis
US-10-237-386-10

Query Match   43.9%;   Score 79;   DB 15;   Length 645;
Best Local Similarity   79.0%;   Pred. No. 2e-17;

```
          STREET: 800 N. Lindbergh
          CITY: St. Louis
          STATE: Missouri
          COUNTRY: USA
          ZIP: 63167
     COMPUTER READABLE FORM:
          MEDIUM TYPE: Floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/10/083,446
          FILING DATE: 26-Feb-2002
          CLASSIFICATION: <Unknown>
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 08/762,227
          FILING DATE: 09-DEC-1996
          APPLICATION NUMBER: US 08/192,325
          FILING DATE: 14-FEB-1994
          APPLICATION NUMBER: US 08/446,872
          FILING DATE: 06-JUN-1995
     ATTORNEY/AGENT INFORMATION:
          NAME: S. Christopher Bauer
          REGISTRATION NUMBER: 42,305
          REFERENCE/DOCKET NUMBER: C-2790/6
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: (636)737-6257
          TELEFAX: (636)737-5452
INFORMATION FOR SEQ ID NO: 177:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 546 base pairs
          TYPE: nucleic acid
          STRANDEDNESS: double
          TOPOLOGY: linear
     MOLECULE TYPE: DNA (genomic)
     SEQUENCE DESCRIPTION: SEQ ID NO: 177:
US-10-083-446-177

Query Match      28.1%;    Score 50.6;   DB 15;   Length 546;
Best Local Similarity  93.0%;  Pred. No. 2.4e-07;
Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy  124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
            ||||| ||| | ||||||||||||||||||||||||||||||||||||||||||||
Db    7 ACACCATTAGGACCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 6
US-10-083-446-178
; Sequence 178, Application US/10083446
; Publication No. US2003018579A1
; GENERAL INFORMATION:
     APPLICANT: Abrams, Mark A.
                Bauer, S. C.
                Braford-Goldberg, Sarah R.
                Caparon, Maire H.
                Easton, Alan M.
                Klein, Barbara K.
                McKearn, John P.
                Olins, Peter O.
                Paik, Kumnan
                Thomas, John W.
     TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
                Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
     NUMBER OF SEQUENCES: 197
     CORRESPONDENCE ADDRESS:
          ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
                Corporate Patent Dept., Mail Zone O4E
          STREET: 800 N. Lindbergh
          CITY: St. Louis
          STATE: Missouri
          COUNTRY: USA
          ZIP: 63167
```

```
     COMPUTER READABLE FORM:
          MEDIUM TYPE: Floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/10/083,446
          FILING DATE: 26-Feb-2002
          CLASSIFICATION: <Unknown>
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 08/762,227
          FILING DATE: 09-DEC-1996
          APPLICATION NUMBER: US 08/192,325
          FILING DATE: 14-FEB-1994
          APPLICATION NUMBER: US 08/446,872
          FILING DATE: 06-JUN-1995
     ATTORNEY/AGENT INFORMATION:
          NAME: S. Christopher Bauer
          REGISTRATION NUMBER: 42,305
          REFERENCE/DOCKET NUMBER: C-2790/6
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: (636)737-6257
          TELEFAX: (636)737-5452
INFORMATION FOR SEQ ID NO: 178:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 546 base pairs
          TYPE: nucleic acid
          STRANDEDNESS: double
          TOPOLOGY: linear
     MOLECULE TYPE: DNA (genomic)
     SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-10-083-446-178

Query Match      28.1%;    Score 50.6;   DB 15;   Length 546;
Best Local Similarity  93.0%;  Pred. No. 2.4e-07;
Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy  124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
            ||||| ||| | ||||||||||||||||||||||||||||||||||||||||||||
Db    7 ACACCATTAGGACCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 7
US-10-083-446-53
; Sequence 53, Application US/10083446
; Publication No. US2003018579A1
; GENERAL INFORMATION:
     APPLICANT: Abrams, Mark A.
                Bauer, S. C.
                Braford-Goldberg, Sarah R.
                Caparon, Maire H.
                Easton, Alan M.
                Klein, Barbara K.
                McKearn, John P.
                Olins, Peter O.
                Paik, Kumnan
                Thomas, John W.
     TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
                Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
     NUMBER OF SEQUENCES: 197
     CORRESPONDENCE ADDRESS:
          ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
                Corporate Patent Dept., Mail Zone O4E
          STREET: 800 N. Lindbergh
          CITY: St. Louis
          STATE: Missouri
          COUNTRY: USA
          ZIP: 63167
     COMPUTER READABLE FORM:
          MEDIUM TYPE: Floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
          CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/10/083,446
            FILING DATE: 26-Feb-2002
            CLASSIFICATION: <Unknown>
          PRIOR APPLICATION DATA:
            APPLICATION NUMBER: 08/762,227
            FILING DATE: 09-DEC-1996
            APPLICATION NUMBER: US 08/192,325
            FILING DATE: 14-FEB-1994
            APPLICATION NUMBER: US 08/446,872
            FILING DATE: 06-JUN-1995
          ATTORNEY/AGENT INFORMATION:
            NAME: S. Christopher Bauer
            REGISTRATION NUMBER: 42,305
            REFERENCE/DOCKET NUMBER: C-2790/6
          TELECOMMUNICATION INFORMATION:
            TELEPHONE: (636)737-6257
            TELEFAX: (636)737-5452
          INFORMATION FOR SEQ ID NO: 53:
            SEQUENCE CHARACTERISTICS:
              LENGTH: 906 base pairs
              TYPE: nucleic acid
              STRANDEDNESS: double
              TOPOLOGY: linear
            MOLECULE TYPE: DNA (genomic)
            SEQUENCE DESCRIPTION: SEQ ID NO: 53:
          US-10-083-446-53

            Query Match     28.1%;   Score 50.6;  DB 15;  Length 906;
            Best Local Similarity    93.0%;  Pred. No. 3e-07;
            Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

          Qy  124 ACTCCGTTAGGTCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
                  || ||||||||| ||| |||||||||||||||||||||||||||||||||||||||
          Db  400 ACACCATTAGGCCCTGCCAGCTCCCTGCCAGCTCCCTGCTCAAGTGCTTAGAG 456

          RESULT 8
          US-10-083-446-72
          ; Sequence 72, Application US/10083446
          ; Publication No. US20030185790A1
          GENERAL INFORMATION:
            APPLICANT: Abrams, Mark A.
                       Bauer, S. C.
                       Braford-Goldberg, Sarah R.
                       Caparon, Maire H.
                       Easton, Alan M.
                       Klein, Barbara K.
                       McKearn, John P.
                       Olins, Peter O.
                       Paik, Kumnan
                       Thomas, John W.
            TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
                                Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
            NUMBER OF SEQUENCES: 197
            CORRESPONDENCE ADDRESS:
              ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
                         Corporate Patent Dept., Mail Zone O4E
              STREET: 800 N. Lindbergh
              CITY: St. Louis
              STATE: Missouri
              COUNTRY: USA
              ZIP: 63167
            COMPUTER READABLE FORM:
              MEDIUM TYPE: Floppy disk
              COMPUTER: IBM PC compatible
              OPERATING SYSTEM: PC-DOS/MS-DOS
              SOFTWARE: PatentIn Release #1.0, Version #1.25
            CURRENT APPLICATION DATA:
              APPLICATION NUMBER: US/10/083,446
              FILING DATE: 26-Feb-2002
              CLASSIFICATION: <Unknown>
            PRIOR APPLICATION DATA:
```

```
          APPLICATION NUMBER: 08/762,227
          FILING DATE: 09-DEC-1996
          APPLICATION NUMBER: US 08/192,325
          FILING DATE: 14-FEB-1994
          APPLICATION NUMBER: US 08/446,872
          FILING DATE: 06-JUN-1995
        ATTORNEY/AGENT INFORMATION:
          NAME: S. Christopher Bauer
          REGISTRATION NUMBER: 42,305
          REFERENCE/DOCKET NUMBER: C-2790/6
        TELECOMMUNICATION INFORMATION:
          TELEPHONE: (636)737-6257
          TELEFAX: (636)737-5452
        INFORMATION FOR SEQ ID NO: 72:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 921 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: double
            TOPOLOGY: linear
          MOLECULE TYPE: DNA (genomic)
          SEQUENCE DESCRIPTION: SEQ ID NO: 72:
        US-10-083-446-72

          Query Match     28.1%;   Score 50.6;  DB 15;  Length 921;
          Best Local Similarity    93.0%;  Pred. No. 3.1e-07;
          Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

        Qy  124 ACTCCGTTAGGTCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
                || ||||||||| ||| |||||||||||||||||||||||||||||||||||||||
        Db    7 ACACCATTAGGCCCTGCCAGCTCCCTGCCAGCTCCCTGCTCAAGTGCTTAGAG 63

        RESULT 9
        US-10-083-446-75
        ; Sequence 75, Application US/10083446
        ; Publication No. US20030185790A1
        GENERAL INFORMATION:
          APPLICANT: Abrams, Mark A.
                     Bauer, S. C.
                     Braford-Goldberg, Sarah R.
                     Caparon, Maire H.
                     Easton, Alan M.
                     Klein, Barbara K.
                     McKearn, John P.
                     Olins, Peter O.
                     Paik, Kumnan
                     Thomas, John W.
          TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
                              Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
          NUMBER OF SEQUENCES: 197
          CORRESPONDENCE ADDRESS:
            ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
                       Corporate Patent Dept., Mail Zone O4E
            STREET: 800 N. Lindbergh
            CITY: St. Louis
            STATE: Missouri
            COUNTRY: USA
            ZIP: 63167
          COMPUTER READABLE FORM:
            MEDIUM TYPE: Floppy disk
            COMPUTER: IBM PC compatible
            OPERATING SYSTEM: PC-DOS/MS-DOS
            SOFTWARE: PatentIn Release #1.0, Version #1.25
          CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/10/083,446
            FILING DATE: 26-Feb-2002
            CLASSIFICATION: <Unknown>
          PRIOR APPLICATION DATA:
            APPLICATION NUMBER: 08/762,227
            FILING DATE: 09-DEC-1996
            APPLICATION NUMBER: US 08/192,325
            FILING DATE: 14-FEB-1994
            APPLICATION NUMBER: US 08/446,872
```

```
        FILING DATE: 06-JUN-1995
      ATTORNEY/AGENT INFORMATION:
        NAME: S. Christopher Bauer
        REGISTRATION NUMBER: 42,305
        REFERENCE/DOCKET NUMBER: C-2790/6
      TELECOMMUNICATION INFORMATION:
        TELEPHONE: (636)737-6257
        TELEFAX: (636)737-5452
INFORMATION FOR SEQ ID NO: 75:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 921 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: double
        TOPOLOGY: linear
      MOLECULE TYPE: DNA (genomic)
      SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-10-083-446-75

Query Match     28.1%;   Score 50.6;  DB 15;  Length 921;
Best Local Similarity  93.0%;  Pred. No. 3.1e-07;
Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy  124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
        || |||||||||| || ||||||||||||||||||||||||||||||||||||||||
Db    7 ACACCATTAGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 10
US-10-083-446-78
; Sequence 78, Application US/10083446
; Publication No. US20030185790A1
GENERAL INFORMATION:
  APPLICANT: Abrams, Mark A.
             Bauer, S. C.
             Braford-Goldberg, Sarah R.
             Caparon, Maire H.
             Easton, Alan M.
             Klein, Barbara K.
             McKearn, John P.
             Olins, Peter O.
             Paik, Kumnan
             Thomas, John W.
  TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
                      Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
  NUMBER OF SEQUENCES: 197
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
               Corporate Patent Dept., Mail Zone O4E
    STREET: 800 N. Lindbergh
    CITY: St. Louis
    STATE: Missouri
    COUNTRY: USA
    ZIP: 63167
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/10/083,446
    FILING DATE: 26-Feb-2002
    CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/762,227
    FILING DATE: 09-DEC-1996
    APPLICATION NUMBER: US 08/192,325
    FILING DATE: 14-FEB-1994
    APPLICATION NUMBER: US 08/446,872
    FILING DATE: 06-JUN-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: S. Christopher Bauer
    REGISTRATION NUMBER: 42,305
    REFERENCE/DOCKET NUMBER: C-2790/6.

      TELECOMMUNICATION INFORMATION:
        TELEPHONE: (636)737-6257
        TELEFAX: (636)737-5452
INFORMATION FOR SEQ ID NO: 78:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 921 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: double
        TOPOLOGY: linear
      MOLECULE TYPE: DNA (genomic)
      SEQUENCE DESCRIPTION: SEQ ID NO: 78:
US-10-083-446-78

Query Match     28.1%;   Score 50.6;  DB 15;  Length 921;
Best Local Similarity  93.0%;  Pred. No. 3.1e-07;
Matches  53;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy  124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
        || |||||||||| || ||||||||||||||||||||||||||||||||||||||||
Db    7 ACACCATTAGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

RESULT 11
US-10-083-446-84
; Sequence 84, Application US/10083446
; Publication No. US20030185790A1
GENERAL INFORMATION:
  APPLICANT: Abrams, Mark A.
             Bauer, S. C.
             Braford-Goldberg, Sarah R.
             Caparon, Maire H.
             Easton, Alan M.
             Klein, Barbara K.
             McKearn, John P.
             Olins, Peter O.
             Paik, Kumnan
             Thomas, John W.
  TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
                      Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
  NUMBER OF SEQUENCES: 197
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
               Corporate Patent Dept., Mail Zone O4E
    STREET: 800 N. Lindbergh
    CITY: St. Louis
    STATE: Missouri
    COUNTRY: USA
    ZIP: 63167
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/10/083,446
    FILING DATE: 26-Feb-2002
    CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/762,227
    FILING DATE: 09-DEC-1996
    APPLICATION NUMBER: US 08/192,325
    FILING DATE: 14-FEB-1994
    APPLICATION NUMBER: US 08/446,872
    FILING DATE: 06-JUN-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: S. Christopher Bauer
    REGISTRATION NUMBER: 42,305
    REFERENCE/DOCKET NUMBER: C-2790/6
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (636)737-6257
    TELEFAX: (636)737-5452
INFORMATION FOR SEQ ID NO: 84:
      SEQUENCE CHARACTERISTICS:
```

;
;   LENGTH: 921 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-10-083-446-84

Query Match   28.1%;   Score 50.6;   DB 15;   Length 921;
Best Local Similarity   93.0%;   Pred. No. 3.1e-07;
Matches   53; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   124 ACTCCGTTAGGTCCAGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
           |||||| |||||| |||||||||||||||||| |||||||||||||||||||||||
Db   400 ACACCATTAGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 456

RESULT 12
US-10-083-446-68
;   Sequence 68, Application US/10083446
;   Publication No. US20030185790A1
;   GENERAL INFORMATION:
;   APPLICANT: Abrams, Mark A.
;   Bauer, S. C.
;   Braford-Goldberg, Sarah R.
;   Caparon, Maire H.
;   Easton, Alan M.
;   Klein, Barbara K.
;   McKearn, John P.
;   Olins, Peter O.
;   Paik, Kumnan
;   Thomas, John W.
;   TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
;   Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
;   NUMBER OF SEQUENCES: 197
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
;   Corporate Patent Dept., Mail Zone O4E
;   STREET: 800 N. Lindbergh
;   CITY: St. Louis
;   STATE: Missouri
;   COUNTRY: USA
;   ZIP: 63167
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/083,446
;   FILING DATE: 26-Feb-2002
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/762,227
;   FILING DATE: 09-DEC-1996
;   APPLICATION NUMBER: US 08/192,325
;   FILING DATE: 14-FEB-1994
;   APPLICATION NUMBER: US 08/446,872
;   FILING DATE: 06-JUN-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: S. Christopher Bauer
;   REGISTRATION NUMBER: 42,305
;   REFERENCE/DOCKET NUMBER: C-2790/6
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (636)737-6257
;   TELEFAX: (636)737-5452
;   INFORMATION FOR SEQ ID NO: 68:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 966 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-083-446-68

Query Match   28.1%;   Score 50.6;   DB 15;   Length 966;
Best Local Similarity   93.0%;   Pred. No. 3.1e-07;
Matches   53; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   124 ACTCCGTTAGGTCCAGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
           |||||| |||||| |||||||||||||||||| |||||||||||||||||||||||
Db   445 ACACCATTAGGCCCTGCCAGCTCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 501

RESULT 13
US-10-083-446-70
;   Sequence 70, Application US/10083446
;   Publication No. US20030185790A1
;   GENERAL INFORMATION:
;   APPLICANT: Abrams, Mark A.
;   Bauer, S. C.
;   Braford-Goldberg, Sarah R.
;   Caparon, Maire H.
;   Easton, Alan M.
;   Klein, Barbara K.
;   McKearn, John P.
;   Olins, Peter O.
;   Paik, Kumnan
;   Thomas, John W.
;   TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
;   Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
;   NUMBER OF SEQUENCES: 197
;   CORRESPONDENCE ADDRESS:
;   ADDRESSE: S. Christopher Bauer, Pharmacia Corporation
;   Corporate Patent Dept., Mail Zone O4E
;   STREET: 800 N. Lindbergh
;   CITY: St. Louis
;   STATE: Missouri
;   COUNTRY: USA
;   ZIP: 63167
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/083,446
;   FILING DATE: 26-Feb-2002
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/762,227
;   FILING DATE: 09-DEC-1996
;   APPLICATION NUMBER: US 08/192,325
;   FILING DATE: 14-FEB-1994
;   APPLICATION NUMBER: US 08/446,872
;   FILING DATE: 06-JUN-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: S. Christopher Bauer
;   REGISTRATION NUMBER: 42,305
;   REFERENCE/DOCKET NUMBER: C-2790/6
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (636)737-6257
;   TELEFAX: (636)737-5452
;   INFORMATION FOR SEQ ID NO: 70:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 966 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-10-083-446-70

Query Match   28.1%;   Score 50.6;   DB 15;   Length 966;
Best Local Similarity   93.0%;   Pred. No. 3.1e-07;

RESULT 15
US-10-083-446-73
; Sequence 73, Application US/10083446
; Publication No. US20030185790A1
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
;   Bauer, S. C.
;   Braford-Goldberg, Sarah R.
;   Caparon, Maire H.
;   Easton, Alan M.
;   Klein, Barbara K.
;   McKearn, John P.
;   Olins, Peter O.
;   Paik, Kumnan
;   Thomas, John W.
; TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
;   Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
;   Corporate Patent Dept., Mail Zone O4E
;   STREET: 800 N. Lindbergh
;   CITY: St. Louis
;   STATE: Missouri
;   COUNTRY: USA
;   ZIP: 63167
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/083,446
;   FILING DATE: 26-Feb-2002
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/762,227
;   FILING DATE: 09-DEC-1996
;   APPLICATION NUMBER: US 08/192,325
;   FILING DATE: 14-FEB-1994
;   APPLICATION NUMBER: US 08/446,872
;   FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
;   NAME: S. Christopher Bauer
;   REGISTRATION NUMBER: 42,305
;   REFERENCE/DOCKET NUMBER: C-2790/6
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (636)737-6257
;   TELEFAX: (636)737-5452
; INFORMATION FOR SEQ ID NO: 73:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 966 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 73:
US-10-083-446-73

Query Match   28.1%; Score 50.6; DB 15; Length 966;
Best Local Similarity   93.0%; Pred. No. 3.1e-07;
Matches   53; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
            || ||| ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   7 ACACCATTAGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 63

Search completed: October 9, 2004, 06:22:17
Job time : 154.651 secs

Matches   53; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
            || ||| ||||| |||||||||||||||||||||||||||||||||||||||||||
Db   445 ACACCATTAGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 501

RESULT 14
US-10-083-446-71
; Sequence 71, Application US/10083446
; Publication No. US20030185790A1
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
;   Bauer, S. C.
;   Braford-Goldberg, Sarah R.
;   Caparon, Maire H.
;   Easton, Alan M.
;   Klein, Barbara K.
;   McKearn, John P.
;   Olins, Peter O.
;   Paik, Kumnan
;   Thomas, John W.
; TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells
;   Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation
;   Corporate Patent Dept., Mail Zone O4E
;   STREET: 800 N. Lindbergh
;   CITY: St. Louis
;   STATE: Missouri
;   COUNTRY: USA
;   ZIP: 63167
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/083,446
;   FILING DATE: 26-Feb-2002
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/762,227
;   FILING DATE: 09-DEC-1996
;   APPLICATION NUMBER: US 08/192,325
;   FILING DATE: 14-FEB-1994
;   APPLICATION NUMBER: US 08/446,872
;   FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
;   NAME: S. Christopher Bauer
;   REGISTRATION NUMBER: 42,305
;   REFERENCE/DOCKET NUMBER: C-2790/6
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (636)737-6257
;   TELEFAX: (636)737-5452
; INFORMATION FOR SEQ ID NO: 71:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 966 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 71:
US-10-083-446-71

Query Match   28.1%; Score 50.6; DB 15; Length 966;
Best Local Similarity   93.0%; Pred. No. 3.1e-07;
Matches   53; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   124 ACTCCGTTAGGTCCAGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 180
            || ||| ||||| |||||||||||||||||||||||||||||||||||||||||||
Db   445 ACACCATTAGGCCCTGCCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAG 501

This Page Blank (uspto)